Robust MIMO Cognitive Radio via Game Theory

Jiaheng Wang, *Member, IEEE*, Gesualdo Scutari, *Member, IEEE*, and Daniel P. Palomar, *Senior Member, IEEE*

***Abstract*—Cognitive Radio (CR) systems improve the spectral efficiency by allowing the coexistence in harmony of primary users (PUs), the legacy users, with secondary users (SUs). This coexistence is built on the premises that no SU can generate interference higher than some prescribed limits against PUs. The system design based on perfect channel state information (CSI) can easily end up violating the interference limits in a realistic situation where CSI may be imperfect. In this paper, we propose a robust design of CR systems, composed of multiple PUs and multiple noncooperative SUs, in either SISO frequency-selective channels or more general MIMO channels. We formulate the design of the SU network as a noncooperative game, where the SUs compete with each other over the resources made available by the PUs, by maximizing their own information rates subject to the transmit power and robust interference constraints. Following the philosophy of the worst-case robustness, we take explicitly into account the imperfectness of SU-to-PU CSI by adopting proper interference constraints that are robust with respect to the worst channel errors. Relying on the variational inequality theory, we study the existence and uniqueness properties of the Nash equilibria of the resulting robust games, and devise totally asynchronous and distributed algorithms along with their convergency properties. We also propose efficient numerical methods, based on decomposition techniques, to compute the robust transmit strategy for each SU.**

***Index Terms*—Cognitive radio, convex optimization, imperfect CSI, MIMO, Nash equilibrium, noncooperative game, variational inequality, worst-case robustness.**

I. Introduction

While the ever-increasing demand of wireless service makes the radio spectrum a very scarce and precious resource, it has been recently reported that the existing fixed spectrum assignment policies cause very inefficient spectrum usage [1]. In such a situation, Cognitive Radio (CR) [2] emerged as a promising approach to improve the efficiency of spectrum utilization by allowing intelligent cognitive nodes to access the licensed bandwidth. A CR system is built on a hierarchical structure [3], [4], distinguishing primary users (PUs), who are legacy spectrum holders, from secondary users (SUs), also known as cognitive users, who are allowed to share the licensed spectrum with PUs, provided that they induce limited interference or no interference at primary receivers [1]. Within this context, alternative approaches have been considered in the literature to allow concurrent communications in either single-input single-output (SISO) channels [5] or multiple-input multiple-output (MIMO) channels [6] (see also [4] for a recent tutorial on the topic).

One classical approach to devise CR systems would be using global optimization techniques, under the framework of network utility maximization (NUM) (see, e.g., [7], [8]) or cooperative games (Nash bargaining solutions) [9], [10]. However, recent results in [11] have shown that NUM problems are NP-hard under different choices of the system utility function. Moreover, the (suboptimal) algorithms proposed in the aforementioned papers lack any mechanism to control the amount of aggregate interference generated by the transmitters, and they are centralized and computationally expensive. This raises some practical issues that are insurmountable in the CR context. For example, these algorithms need a central node having full knowledge of all the channels and interference structure at every receiver, which poses serious implementation problems in terms of scalability and amount of signaling to be exchanged among the nodes. For these reasons, in this paper, we follow a different approach and model the SUs as competitors with no will to cooperate. We concentrate on *decentralized* strategies, where the SUs are able to self-enforce the negotiated agreements on the usage of the available spectrum without the intervention of a centralized authority. This form of equilibrium is, in fact, the well-known concept of Nash equilibrium (NE) in noncooperative game theory [12]. In the literature, noncooperative game theory has been successfully applied to the design of the network over SISO frequency-selective interference channels [13]–[18] as well as MIMO interference channels [19], [20], and recently to the CR framework [21]–[24].

The aforementioned works were all based on the assumption of perfect channel state information between SUs and PUs (SU-to-PU CSI). In practice, however, SU-to-PU CSI is seldom perfectly known, due to the loose or usually no cooperation between PUs and SUs, as well as many other factors such as inaccurate channel estimation and time delays or frequency offsets between the reciprocal channels. Consequently, although designated to meet the interference constraints, an SU may still break this limitation because of imperfect SU-to-PU CSI. This violation of the interference constraints cannot be tolerated in a CR system and therefore the imperfectness of SU-to-PU CSI has to be taken into account in the system design. There are two common ways to model imperfect CSI: Bayesian and worst-case approaches. The Bayesian philosophy assumes that the channel is a ran-

Copyright (c) 2010 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

J. Wang and D. P. Palomar are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong (e-mail: jhwang@ust.hk, palomar@ust.hk).

G. Scutari is with the Department of Electrical Engineering at State University of New York (SUNY) at Buffalo, Buffalo, NY, USA (email: gesualdo@buffalo.edu).

This work was supported by the Hong Kong RGC 618709 research grant.

VISIT…

dom quantity and guarantees the constraint on the average. Although this model has been adopted to optimize the average performance of conventional communications [25], [26], it may not be suitable for a CR system, because an average interference constraint may not be acceptable to the PU as it may affect the normal communication of the PU. The worst-case approach, more suitable for characterizing instantaneous CSI with errors, assumes that the actual channel lies in the neighborhood–the so-called channel uncertainty region–of a known nominal channel (e.g., the estimated channel). The size of this region represents the amount of channel uncertainty, i.e., the bigger the region is the more uncertainty there is. Then, using the worst-case robust design [27]–[33], one can guarantee a performance level for any channel realization in the uncertainty region. This meets exactly the requirement of a CR system—the interference constraints must be satisfied even if PU CSI is imperfect. Hence, it is not surprising that the idea of the worst-case robustness has recently been used to design robust beamforming for SUs in a multiple-input single-output (MISO) CR system [34], [35].

In this paper, we consider the design of robust CR systems with multiple PUs and multiple noncooperative SUs in either SISO frequency-selective channels or MIMO channels (see Fig. 1). Our goal is to devise the most appropriate form of concurrent communications of the SUs competing over the resources that the PUs make available, while taking into account imperfect SU-to-PU CSI. Given the competitive nature of the SU network, we formulate the resource allocation problem among SUs as a strategic noncooperative game, where each SU competes against the others to maximize his own information rate under the power constraints and the robust interference constraints. Following the worst-case robustness philosophy, the imperfectness of SU-to-PU CSI is explicitly taken into account by the robust interference constraints. The main contributions of the paper are the following.

First, we consider robust issues in game theoretical formulations, which have been seldom investigated in existing works. In particular, if the uncertainty on CSI is independently modeled for each subchannel as in [36], then the game formulation boils down into the problems considered in [14]–[20], [23], [24]. In this paper, we consider a more practical and general situation where the channel uncertainty is coupled across frequency dimensions in SISO frequency-selective channels or spatial dimensions in MIMO channels.

Second, the presence of the robust interference constraints leads to an intractable optimization problem for each player, as the constraint set is given in the form of the intersection of an infinite number of convex sets. We show that this difficulty can be overcome by rewriting the constraint set into an equivalent and more convenient form.

Third, the robust game formulation, even with the equivalent constraint set, is still difficult to study, because there is no closed-form expression for the best-response of each player. This means that all previous results based the fixed-point theory in [16], [17], [19], [20], [23] are no longer applicable. Herein, we rely on the more advanced theory of finite-dimensional variational inequalities (VIs) [37], [38]. We study the existence and uniqueness properties of the NE of the robust games, and propose totally asynchronous and distributed algorithms along with their convergence properties to achieve the NE.

Finally, we propose efficient and low-complexity numerical methods, based on decomposition techniques [8], to compute the best-response–the robust transmit strategy–of each SU. Note that our framework contains all existing single-user robust CR designs (e.g., [34]) as special cases.

The paper is organized as follows. Section II introduces the system model, robust interference constraints, and robust game formulations. In Section III, the robust interference constraint sets are equivalently transformed into the more favorable forms. Then, we study the robust games via the VI theory, and devise distributed algorithms for the SU network in Section IV. In Section V, we show how to obtain the robust transmit strategy for each SU through convex optimization. Numerical results are provided in Section VI; and Section VII draws the conclusion.

Notation: Uppercase and lowercase boldface denote matrices and vectors, respectively. $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{S}_+$ denote the sets of real numbers, complex numbers, and positive semidefinite matrices, respectively. $[\mathbf{X}]_{ij}$ represents the (ith, jth) element of matrix $\mathbf{X}$. By $\mathbf{X} \succeq \mathbf{0}$ or $\mathbf{X} \succ \mathbf{0}$, we mean that $\mathbf{X}$ is a positive semidefinite or definite matrix, respectively. $\mathbf{I}$ and $\mathbf{1}$ denote the identity matrix and the vector of ones, respectively. The operators $\geq$ and $\leq$ for vectors and matrices are defined componentwise. The operators $(\cdot)^H$, $(\cdot)^{-1}$, vec$(\cdot)$, and Tr$(\cdot)$ denote the Hermitian, inverse, stacking vectorization, and trace operations, respectively, The operator diag$(\cdot)$ is the diagonal matrix with the same elements or diagonal elements of a vector or matrix, respectively. The spectral radius [39] of a matrix is denoted by $\rho(\cdot)$, and the maximum and minimum eigenvalues of a Hermitian matrix are denoted by $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$, respectively. $\|\cdot\|$ denotes the Euclidean norm of a vector, and $\|\cdot\|_F$ and $\|\cdot\|_2$ denote the Frobenius and spectral norms of a matrix, respectively. $\otimes$ and $\odot$ represent the Kronecker and Hadamard product operators, respectively, and $(x)_+ \triangleq \max(0, x)$.

II. System Model and Problem Statement

We consider a hierarchical CR network composed of K PUs and Q SUs, coexisting in the same area and sharing the same spectrum (see Fig. 1). Because of the lack of coordination among all users, the set of SUs can be naturally modeled as a vector Gaussian interference channel. Our interest lies in: 1) devising the most appropriate form of concurrent communications of SUs competing over the physical resources that PUs make available, while preserving the Quality of Service (QoS) of PUs; and 2) improving the robustness of the CR system against the imperfectness of SU-to-PU CSI.

Due to the distributed and competitive nature of the CR system, we concentrate on decentralized strategies, where SUs are able to self-enforce the negotiated agreements on the usage of the available spectrum without the intervention of a centralized authority. More specifically, adopting an information theoretical perspective, we formulate the resource allocation problem among SUs as a strategic noncooperative

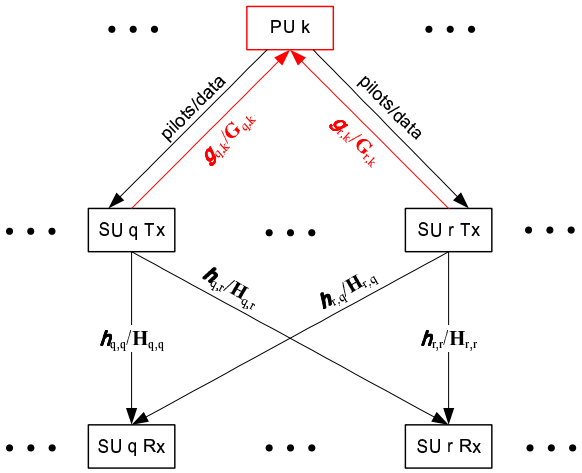


Fig. 1. System model of a hierarchical CR system with multiple PUs and multiple noncooperative SUs.

game, where each player (transmit–receive pair) competes against the others to maximize the information rate over his own channel, under constraints on the transmit power and the interference induced to PUs. Imperfect SU-to-PU CSI is explicitly taken into account by introducing proper interference constraints that are robust with respect to the worst channel errors. We consider transmissions over both SISO frequency-selective channels and MIMO channels.

A. SISO Frequency-Selective Channels

Suppose that the bandwidth shared by the PUs and SUs can be divided into N frequency bins. Let $h_{rq}(n)$ be the cross-channel transfer function over the n^{th} frequency bin between the secondary transmitter r and the secondary receiver q, for $n = 1, \ldots, N$, and $q, r = 1, \ldots, Q$; whereas the channel transfer function of link q over frequency bin n is denoted by $h_{qq}(n)$. The cross-channel transfer function between the transmitter of SU q and the receiver of PU k over the n^{th} frequency bin is denoted by $g_{qk}(n)$, $k = 1, \ldots, K$. The transmit strategy of each SU q is given by his power allocation vector $\mathbf{p}_q \triangleq (p_q(n))_{n=1}^N$ over the N frequency bins, subject to the following power constraints:

$$\mathcal{P}_q^{\text{pow}} \triangleq \left\{ \mathbf{p}_q \in \mathbb{R}^N : \mathbf{1}^T \mathbf{p}_q \le P_q^{\text{ave}},\ \mathbf{0} \le \mathbf{p}_q \le \mathbf{p}_q^{\text{peak}} \right\} \quad (1)$$

where P_q^{ave} is the transmit power in units of energy per transmission, and $\mathbf{p}_q^{\text{peak}} \triangleq (p_q^{\text{peak}}(n))_{n=1}^N$ are the peak power (or spectral mask) limits. Assuming w.l.o.g that each SU as well as each PU uses a Gaussian codebook [40], the maximum achievable rate on SU link q for a given power allocation profile $\mathbf{p} \triangleq (\mathbf{p}_q)_{q=1}^Q$ is

$$r_q(\mathbf{p}_q, \mathbf{p}_{-q}) \triangleq \sum_{n=1}^N \log \left(1 + \frac{|h_{qq}(n)|^2 p_q(n)}{\sigma_q^2(n) + \sum_{r \ne q} |h_{rq}(n)|^2 p_r(n)} \right) \quad (2)$$

where $\mathbf{p}_{-q} \triangleq (\mathbf{p}_r)_{r \ne q}$ denote power allocation of all users except the q^{th} one, and $\sigma_q^2(n)$ includes the noise power plus the interference from all PUs.

Interference constraints. Opportunistic communications in CR systems enable SUs to transmit with overlapping spectrum and/or coverage with PUs, provided that the degradation induced on PUs' performance is null or tolerable [3], [4]. This can be handled by, e.g., introducing some interference constraints that impose upper bounds on the per-carrier and total aggregate interference caused by one SU to one PU. Under the assumption of perfect SU-to-PU CSI, the interference constraints on each SU q can be written as

$$\mathcal{P}_q^{\text{int}} \triangleq \left\{ \mathbf{p}_q \in \mathbb{R}^N : \begin{array}{l} \sum_{n=1}^N |g_{qk}(n)|^2 p_q(n) \le I_{qk}^{\text{ave}}, \\ |g_{qk}(n)|^2 p_q(n) \le I_{qk}^{\text{peak}}(n),\ \forall n, k \end{array} \right\} \quad (3)$$

where I_{qk}^{ave} is the maximum average interference[1] that can be generated by SU q at primary receiver k, and $I_{qk}^{\text{peak}}(n)$ is the maximum peak interference that can be generated over the n^{th} frequency bin.

In practice, however, SU-to-PU CSI is seldom perfect due to many issues. To take into account imperfect SU-to-PU CSI, we adopt the following common imperfect CSI model [27]–[35]: the actual channel is assumed to be within the neighborhood of a nominal channel, while the nominal channel could be the estimated or feedback channel. Specifically, for PU k, SU q only knows a nominal channel $\hat{\boldsymbol{g}}_{qk} \triangleq (\hat{g}_{qk}(n))_{n=1}^N$, which is a corrupted version of the actual channel $\boldsymbol{g}_{qk} \triangleq (g_{qk}(n))_{n=1}^N$ by an error $\boldsymbol{e}_{qk} \triangleq (e_{qk}(n))_{n=1}^N$, i.e.,

$$\hat{\boldsymbol{g}}_{qk} = \boldsymbol{g}_{qk} - \boldsymbol{e}_{qk} \quad (4)$$

where $\boldsymbol{e}_{qk}$ belongs to an elliptical uncertainty region $\mathcal{D}_{qk}$ defined by the weighted Euclidean norm as

$$\begin{aligned} \mathcal{D}_{qk} &\triangleq \left\{ \boldsymbol{e}_{qk} \in \mathbb{C}^N : \|\boldsymbol{e}_{qk}\|_{\mathbf{w}_{qk}} \le \varepsilon_{qk} \right\} \\ &= \left\{ \boldsymbol{e}_{qk} \in \mathbb{C}^N : \sum_{n=1}^N |e_{qk}(n)|^2 w_{qk}(n) \le \varepsilon_{qk}^2 \right\} \end{aligned} \quad (5)$$

where $\mathbf{w}_{qk} \triangleq (w_{qk}(n))_{n=1}^N$ are given positive weights. The radius ε_{qk} represents the size of the uncertainty region, i.e., the larger the radius ε_{qk} is, the more uncertainty there is. The properness of this model has been justified in [29], [30]. In particular, when $\mathbf{w}_{qk} = \mathbf{1}$, the ellipsoid reduces to a sphere, which is the most common model in the literature [27]–[30]. Following the philosophy of worst-case robustness, an SU should keep his interference against a PU below the required thresholds for any possible channel (error) in the uncertainty region. More specifically, each SU q should satisfy, instead of the constraints in (3), the robust interference constraints in (6) at the bottom of the next page.

Game theoretical formulation. The resource allocation problem among the SUs is formulated as a *robust* strategic noncooperative game $\hat{\mathcal{G}}_{\text{siso}} = \langle \Omega, \{\hat{\mathcal{P}}_q\}_{q \in \Omega}, \{r_q\}_{q \in \Omega} \rangle$, where $\Omega \triangleq \{1, \ldots, Q\}$ is the set of the Q SUs, r_q is the utility function defined in (2), and

$$\hat{\mathcal{P}}_q \triangleq \mathcal{P}_q^{\text{pow}} \cap \hat{\mathcal{P}}_q^{\text{int}} \quad (7)$$

is the set of robust admissible strategies for SU q, where $\mathcal{P}_q^{\text{pow}}$ and $\hat{\mathcal{P}}_q^{\text{int}}$ are defined in (1) and (6), respectively. Each SU q's

[1]Note that there is also another kind of average interference constraints, where the average is taken over the random channel, for which the interested reader is referred to [41].

problem is to determine, for a fixed but arbitrary power profile $\mathbf{p}_{-q}$ of the other SUs, an optimal strategy $\mathbf{p}_q^\star$ that solves the following optimization problem in the variable $\mathbf{p}_q$

$$\underset{\mathbf{p}_q \in \hat{\mathcal{P}}_q}{\text{maximize}} \; r_q(\mathbf{p}_q, \mathbf{p}_{-q}). \tag{8}$$

A strategy profile $\mathbf{p}^\star = (\mathbf{p}_q^\star)_{q=1}^Q$ is a NE of $\hat{\mathcal{G}}_{\text{siso}}$ if $r_q(\mathbf{p}_q^\star, \mathbf{p}_{-q}^\star) \geq r_q(\mathbf{p}_q, \mathbf{p}_{-q}^\star)$, $\forall \mathbf{p}_q \in \hat{\mathcal{P}}_q$, $\forall q \in \Omega$. In the forthcoming sections, we provide a full characterization of the robust game $\hat{\mathcal{G}}_{\text{siso}}$. For comparison, we also consider the *non-robust* game $\mathcal{G}_{\text{siso}} = \langle \Omega, \{\mathcal{P}_q\}_{q\in\Omega}, \{r_q\}_{q\in\Omega}\rangle$, with the non-robust strategy set $\mathcal{P}_q \triangleq \mathcal{P}_q^{\text{pow}} \cap \mathcal{P}_q^{\text{int}}$, where each SU q regards the nominal channels $\hat{\boldsymbol{g}}_{qk}$, $\forall k$, as the actual channels.

B. MIMO Channels

In this section, we focus on transmissions over MIMO channels: each SU link q is equipped with N_q and M_q transmit and receive antennas, respectively; the receiver of each PU k is equipped with M_k antennas. Let $\mathbf{H}_{rq} \in \mathbb{C}^{M_q \times N_r}$ be the cross-channel transfer function between the secondary transmitter r and the secondary receiver q; whereas the channel transfer function of link q is denoted by $\mathbf{H}_{qq}$. The cross-channel transfer function between the transmitter of SU q and the receiver of PU k is denoted by $\mathbf{G}_{qk} \in \mathbb{C}^{M_k \times N_q}$. The transmit strategy of SU q is determined by his transmit covariance matrix $\mathbf{Q}_q$, which should satisfy the following power constraints:

$$\mathcal{Q}_q^{\text{pow}} \triangleq \left\{ \mathbf{Q}_q \in \mathbb{S}_+^{N_q} : \text{Tr}(\mathbf{Q}_q) \leq P_q^{\text{ave}}, \; \lambda_{\max}(\mathbf{Q}_q) \leq P_q^{\text{peak}} \right\} \tag{9}$$

where P_q^{ave} is the transmit power in units of energy per transmission, and P_q^{peak} is the spacial peak average power constraint. With each SU and each PU using Gaussian codebooks [40], the maximum achievable rate on SU link q for a given profile $\mathbf{Q} \triangleq (\mathbf{Q}_q)_{q=1}^Q$ is

$$R_q(\mathbf{Q}_q, \mathbf{Q}_{-q}) \triangleq \log\det(\mathbf{I} + \mathbf{H}_{qq}^H \mathbf{R}_{-q}^{-1}(\mathbf{Q}_{-q}) \mathbf{H}_{qq} \mathbf{Q}_q), \tag{10}$$

where $\mathbf{Q}_{-q} \triangleq (\mathbf{Q}_r)_{r\neq q}$, $\mathbf{R}_{-q}(\mathbf{Q}_{-q}) \triangleq \mathbf{R}_{n_q} + \sum_{r\neq q} \mathbf{H}_{rq}\mathbf{Q}_r\mathbf{H}_{rq}^H$, and $\mathbf{R}_{n_q} \succ \mathbf{0}$ is the covariance matrix of the noise plus the interference from the PUs.

Interference constraints. Under the assumption of perfect CSI of PUs, the interference constraints imposed to each SU q in the MIMO case are given by

$$\mathcal{Q}_q^{\text{int}} \triangleq \left\{ \begin{array}{ll} \mathbf{Q}_q \in \mathbb{S}_+^{N_q} : & \text{Tr}(\mathbf{G}_{qk}\mathbf{Q}_q\mathbf{G}_{qk}^H) \leq I_{qk}^{\text{ave}}, \\ & \lambda_{\max}(\mathbf{G}_{qk}\mathbf{Q}_q\mathbf{G}_{qk}^H) \leq I_{qk}^{\text{peak}}, \; \forall k \end{array} \right\} \tag{11}$$

where I_{qk}^{ave} and I_{qk}^{peak} are the maximum average and peak interference that can be generated by SU q at the receiver of PU k, respectively.

To characterize the imperfectness of SU-to-PU CSI, we use a similar model as in the SISO case. More specifically, each secondary transmitter q is assumed to know a corrupted version $\hat{\mathbf{G}}_{qk}$ of the actual channel $\mathbf{G}_{qk}$, given by

$$\hat{\mathbf{G}}_{qk} = \mathbf{G}_{qk} - \mathbf{E}_{qk} \tag{12}$$

where $\mathbf{E}_k$ is an error matrix belonging to an elliptical uncertainty region $\mathcal{E}_{qk}$, defined by the weighted Frobenius norm as

$$\begin{aligned} \mathcal{E}_{qk} &\triangleq \{\mathbf{E}_{qk} \in \mathbb{C}^{M_k \times N_q} : \|\mathbf{E}_{qk}\|_{\mathbf{T}_k,F} \leq \varepsilon_{qk}\} \\ &= \{\mathbf{E}_{qk} \in \mathbb{C}^{M_k \times N_q} : \text{Tr}\left(\mathbf{E}_{qk}\mathbf{T}_{qk}\mathbf{E}_{qk}^H\right) \leq \varepsilon_{qk}^2\} \end{aligned} \tag{13}$$

with the weight matrix $\mathbf{T}_k \succ \mathbf{0}$. Similarly, the ellipsoid reduces to a sphere when $\mathbf{T}_k = \mathbf{I}$, as the most frequently used model [27]–[30]. As the interference limits must be strictly met for any channel (error) in the uncertainty region, SU q should satisfy, instead of the constraints in (11), the robust interference constraints in (14) at the bottom of the page.

Game theoretical formulation. We formulate optimization of the transmit strategies of the SUs as a *robust* strategic noncooperative game $\hat{\mathcal{G}}_{\text{mimo}} = \langle \Omega, \{\hat{\mathcal{Q}}_q\}_{q\in\Omega}, \{R_q\}_{q\in\Omega}\rangle$, where R_q is the utility function defined in (10), and

$$\hat{\mathcal{Q}}_q \triangleq \mathcal{Q}_q^{\text{pow}} \cap \hat{\mathcal{Q}}_q^{\text{int}} \tag{15}$$

is the set of robust admissible strategies for SU q, where $\mathcal{Q}_q^{\text{pow}}$ and $\hat{\mathcal{Q}}_q^{\text{int}}$ are defined in (9) and (14), respectively. Each SU q tries to find his optimal transmit strategy $\mathbf{Q}_q^\star$, given the transmit strategies $\mathbf{Q}_{-q}$ of the others, by solving the following optimization problem:

$$\underset{\mathbf{Q}_q \in \hat{\mathcal{Q}}_q}{\text{maximize}} \; R_q(\mathbf{Q}_q, \mathbf{Q}_{-q}). \tag{16}$$

We will study the robust game $\hat{\mathcal{G}}_{\text{mimo}}$ in the following sections. For comparison, we also consider the *non-robust* game $\mathcal{G}_{\text{mimo}} = \langle \Omega, \{\mathcal{Q}_q\}_{q\in\Omega}, \{R_q\}_{q\in\Omega}\rangle$ with the non-robust strategy set $\mathcal{Q}_q \triangleq \mathcal{Q}_q^{\text{pow}} \cap \mathcal{Q}_q^{\text{int}}$, where each SU q regards the nominal channels $\hat{\mathbf{G}}_{qk}$, $\forall k$, as the actual channels.

III. Characterization of the Robust Interference Constraint Sets

The presence of the robust interference constraints makes the analysis of the games $\hat{\mathcal{G}}_{\text{siso}}$ and $\hat{\mathcal{G}}_{\text{mimo}}$ quite involved. The reason is that the robust interference constraint sets $\hat{\mathcal{P}}_q^{\text{int}}$

$$\hat{\mathcal{P}}_q^{\text{int}} \triangleq \left\{ \begin{array}{ll} \mathbf{p}_q \in \mathbb{R}^N : & \sum_{n=1}^N |\hat{g}_{qk}(n) + e_{qk}(n)|^2 p_q(n) \leq I_{qk}^{\text{ave}}, \\ & |\hat{g}_{qk}(n) + e_{qk}(n)|^2 p_q(n) \leq I_{qk}^{\text{peak}}(n), \; \forall n, \forall \boldsymbol{e}_{qk} \in \mathcal{D}_{qk}, \forall k \end{array} \right\} \tag{6}$$

$$\hat{\mathcal{Q}}_q^{\text{int}} \triangleq \left\{ \begin{array}{ll} \mathbf{Q}_q \in \mathbb{S}_+^{N_q} : & \text{Tr}((\hat{\mathbf{G}}_{qk} + \mathbf{E}_{qk})\mathbf{Q}_q(\hat{\mathbf{G}}_{qk} + \mathbf{E}_{q,k})^H) \leq I_{qk}^{\text{ave}}, \\ & \lambda_{\max}((\hat{\mathbf{G}}_{qk} + \mathbf{E}_{qk})\mathbf{Q}_q(\hat{\mathbf{G}}_{qk} + \mathbf{E}_{qk})^H) \leq I_{qk}^{\text{peak}}, \; \forall \mathbf{E}_{qk} \in \mathcal{E}_{qk}, \forall k \end{array} \right\} \tag{14}$$

and $\hat{\mathcal{Q}}_q^{\rm int}$ are given in the form of the intersection of an infinite number of convex sets. In this section, we overcome this difficulty by rewriting the sets $\hat{\mathcal{P}}_q^{\rm int}$ and $\hat{\mathcal{Q}}_q^{\rm int}$ into more convenient and equivalent forms, which paves the way to studying the robust games $\hat{\mathcal{G}}_{\rm siso}$ and $\hat{\mathcal{G}}_{\rm mimo}$, and designing distributed algorithms along with their convergence properties.

A. SISO Frequency-Selective Channels

The set $\hat{\mathcal{P}}_q^{\rm int}$ in (6) can be rewritten in a more convenient form as given next.

Proposition 1: Given $\boldsymbol{\varepsilon} \triangleq \{\{\varepsilon_{qk}\}_{k=1}^K\}_{q=1}^Q > \mathbf{0}$, the set $\hat{\mathcal{P}}_q^{\rm int}$ is equivalent to (17) at the bottom of the page.

Proof: We first introduce the following lemma.

Lemma 1 (S-procedure [42]): Let $f_k(\mathbf{x})$, $k = 1, 2$, be defined as

$$f_k(\mathbf{x}) = \mathbf{x}^H \mathbf{A}_k \mathbf{x} + 2\mathrm{Re}\left\{\mathbf{b}_k^H \mathbf{x}\right\} + c_k$$

where $\mathbf{A}_k = \mathbf{A}_k^H \in \mathbb{C}^{n\times n}$, $\mathbf{b}_k \in \mathbb{C}^n$ and $c_k \in \mathbb{R}$. Then, the implication $f_1(\mathbf{x}) \geq 0 \Rightarrow f_2(\mathbf{x}) \geq 0$ holds if and only if there exists $\mu \geq 0$ such that

$$\begin{bmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^H & c_2 \end{bmatrix} - \mu \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^H & c_1 \end{bmatrix} \succeq \mathbf{0},$$

provided that there exists a point $\hat{\mathbf{x}}$ with $f_1(\hat{\mathbf{x}}) > 0$. □

Consider first the robust sum interference constraint in $\hat{\mathcal{P}}_q^{\rm int}$

$$\sum_{n=1}^N |\hat{g}_{qk}(n) + e_{qk}(n)|^2 p_q(n) \leq I_{qk}^{\rm ave},\ \forall \boldsymbol{e}_{qk} \in \mathcal{D}_{qk} \tag{18}$$

which, by introducing $\mathbf{P}_q \triangleq \mathrm{diag}\{p_q(n)\}_{n=1}^N$ and $\mathbf{W}_{qk} \triangleq \mathrm{diag}\{w_{qk}(n)\}_{n=1}^N$, can be expressed as

$$\begin{aligned} -\boldsymbol{e}_{qk}^H \mathbf{P}_q \boldsymbol{e}_{qk} - 2\mathrm{Re}\{\hat{\boldsymbol{g}}_{qk}^H \mathbf{P}_q \boldsymbol{e}_{qk}\} - \hat{\boldsymbol{g}}_{qk}^H \mathbf{P}_q \hat{\boldsymbol{g}}_{qk} + I_{qk}^{\rm ave} \geq 0, \\ \forall \boldsymbol{e}_{qk} : -\boldsymbol{e}_{qk}^H \mathbf{W}_{qk} \boldsymbol{e}_{qk} + \varepsilon_{qk}^2 \geq 0. \end{aligned} \tag{19}$$

According to Lemma 1, (19) holds if and only if there exists $\mu_{qk} \geq 0$ such that

$$\begin{bmatrix} \mu_{qk}\mathbf{W}_{qk} - \mathbf{P}_q & -\mathbf{P}_q \hat{\boldsymbol{g}}_{qk} \\ -\hat{\boldsymbol{g}}_{qk}^H \mathbf{P}_q & I_{qk}^{\rm ave} - \mu_{qk}\varepsilon_{qk}^2 - \hat{\boldsymbol{g}}_{qk}^H \mathbf{P}_q \hat{\boldsymbol{g}}_{qk} \end{bmatrix} \succeq \mathbf{0}. \tag{20}$$

Note that the condition $f_1(\hat{\mathbf{x}}) > 0$ is satisfied since $\varepsilon_{qk}^2 > 0$.

Next, consider the robust peak interference constraint

$$|\hat{g}_{qk}(n) + e_{qk}(n)|^2 p_q(n) \leq I_{qk}^{\rm peak}(n),\ \forall n,\ \forall \boldsymbol{e}_{qk} \in \mathcal{D}_{qk} \tag{21}$$

which can be rewritten as

$$\max_{\|\tilde{\mathbf{e}}_{qk}\| \leq \varepsilon_{qk}} \max_n \left\{|\tilde{g}_{qk}(n) + \tilde{e}_{qk}(n)|^2 \tilde{p}_{qk}(n)\right\} \leq 1 \tag{22}$$

where $\tilde{\boldsymbol{e}}_{qk} \triangleq \mathbf{W}_{qk}^{1/2} \boldsymbol{e}_{qk}$, $\tilde{g}_{qk}(n) \triangleq \sqrt{w_{qk}(n)}\hat{g}_{qk}(n)$, and $\tilde{p}_{qk}(n) \triangleq p_q(n)/(w_{qk}(n) I_{qk}^{\rm peak}(n))$. It follows that

$$\begin{aligned} &\max_{\|\tilde{\mathbf{e}}_{qk}\| \leq \varepsilon_{qk}} \max_n \left\{|\tilde{g}_{qk}(n) + \tilde{e}_{qk}(n)|^2 \tilde{p}_{qk}(n)\right\} \\ &\quad = \max_n \max_{\|\tilde{\mathbf{e}}_{qk}\| \leq \varepsilon_{qk}} \left\{|\tilde{g}_{qk}(n) + \tilde{e}_{qk}(n)|^2 \tilde{p}_{qk}(n)\right\} \\ &\quad = \max_n \left\{(|\tilde{g}_{qk}(n)| + \varepsilon_{qk})^2 \tilde{p}_{qk}(n)\right\}. \end{aligned} \tag{23}$$

Hence, (22) amounts to

$$(|\tilde{g}_{qk}(n)| + \varepsilon_{qk})^2 \tilde{p}_{qk}(n) \leq 1,\ \forall n \tag{24}$$

or equivalently

$$\left(|\hat{g}_{qk}(n)| + \varepsilon_{qk}/\sqrt{w_{qk}(n)}\right)^2 p_q(n) \leq I_{qk}^{\rm peak}(n),\ \forall n. \tag{25}$$

This completes the proof. ■

The interference constraints in the equivalent set (17) are linear matrix inequalities (LMIs), which are a generalization of linear inequalities in $\mathbb{S}_+$ [42]. Therefore, (17) is in fact an intersection of the sublevels[2] of linear functions, and thus much more convenient than the original form (6). Furthermore, many numerical methods, for example the interior-point method [43], require the constraint set of a convex optimization problem to be an intersection of sublevels of some convex functions. Therefore, Proposition 1 not only provides a simpler strategy set to study the game, but also is an indispensable step to efficiently compute the solution to each single-user optimization problem in (8) (the best-response of each player).

[2]An α-sublevel of a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is defined as $\{x \in \mathrm{dom} f : f(x) \leq \alpha\}$.

$$\hat{\mathcal{P}}_q^{\rm int} = \left\{ \begin{array}{ll} \mathbf{p}_q \in \mathbb{R}^N : & \exists \boldsymbol{\mu}_q \triangleq (\mu_{qk})_{k=1}^K \geq \mathbf{0}, \\ & \begin{bmatrix} \mathrm{diag}\{\mu_{qk}\mathbf{w}_{qk} - \mathbf{p}_q\} & -\mathbf{p}_q \odot \hat{\boldsymbol{g}}_{qk} \\ -(\mathbf{p}_q \odot \hat{\boldsymbol{g}}_{qk})^H & I_{qk}^{\rm ave} - \mu_{qk}\varepsilon_{qk}^2 - \hat{\boldsymbol{g}}_{qk}^H \mathrm{diag}\{\mathbf{p}_q\}\hat{\boldsymbol{g}}_{qk} \end{bmatrix} \succeq \mathbf{0}, \\ & \left(|\hat{g}_{qk}(n)| + \varepsilon_{qk}/\sqrt{w_{qk}(n)}\right)^2 p_q(n) \leq I_{qk}^{\rm peak}(n),\ \forall n, k \end{array} \right\} \tag{17}$$

$$\hat{\mathcal{Q}}_q^{\rm int} = \left\{ \begin{array}{ll} \mathbf{Q}_q \in \mathbb{S}_+^{N_q} : & \exists \boldsymbol{\mu}_q \triangleq (\mu_{qk})_{k=1}^K \geq \mathbf{0},\ \exists \boldsymbol{\eta}_q \triangleq (\eta_{qk})_{k=1}^K \geq \mathbf{0}, \\ & \begin{bmatrix} (\mu_{qk}\mathbf{T}_{qk} - \mathbf{Q}_q)^T \otimes \mathbf{I} & -\mathrm{vec}(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q) \\ -\mathrm{vec}^H(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q) & I_{qk}^{\rm ave} - \mu_{qk}\varepsilon_{qk}^2 - \mathrm{Tr}(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H) \end{bmatrix} \succeq \mathbf{0}, \\ & \begin{bmatrix} I_{qk}^{\rm peak}\mathbf{I} - \eta_{qk}\varepsilon_{qk}^2\mathbf{I} - \hat{\mathbf{G}}_{qk}\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H & -\hat{\mathbf{G}}_{qk}\mathbf{Q}_q \\ -\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H & \eta_{qk}\mathbf{T}_{qk} - \mathbf{Q}_q \end{bmatrix} \succeq \mathbf{0},\ \forall k \end{array} \right\} \tag{26}$$

B. MIMO Channels

In the case of MIMO channels, the set $\hat{\mathcal{Q}}_q^{\rm int}$ in (14) can be equivalently rewritten as follows.

Proposition 2: Given $\boldsymbol{\varepsilon} \triangleq \{\{\varepsilon_{qk}\}_{k=1}^K\}_{q=1}^Q > \mathbf{0}$, the set $\hat{\mathcal{Q}}_q^{\rm int}$ is equivalent to (26) at the bottom of the previous page.

Proof: The proof is based on Lemma 1 and the following intermediate result.

Lemma 2 ([44, Proposition 3.4]): Provided $\mathbf{D} \succeq \mathbf{0}$, the condition

$$\mathbf{C} + \mathbf{B}\mathbf{X}^H + \mathbf{X}\mathbf{B}^H + \mathbf{X}\mathbf{A}\mathbf{X}^H \succeq \mathbf{0}, \quad \forall \mathbf{X} : \mathrm{Tr}(\mathbf{X}\mathbf{D}\mathbf{X}^H) \leq 1$$

holds if and only if there exists $\omega \geq 0$ such that

$$\left[\begin{array}{cc} \mathbf{C} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{A} \end{array} \right] - \omega \left[\begin{array}{cc} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{D} \end{array} \right] \succeq \mathbf{0}.$$

□

First, consider in $\hat{\mathcal{Q}}_q^{\rm int}$ the constraint

$$\mathrm{Tr}((\hat{\mathbf{G}}_{qk}+\mathbf{E}_{qk})\mathbf{Q}_q(\hat{\mathbf{G}}_{qk}+\mathbf{E}_{qk})^H) \leq I_{qk}^{\rm ave}, \ \forall \mathbf{E}_{qk} \in \mathcal{E}_{qk} \quad (27)$$

which is equal to

$$\begin{aligned} \mathbf{e}_{qk}^H(\mathbf{Q}_q^T \otimes \mathbf{I})\mathbf{e}_{qk} - 2\mathrm{Re}\{\mathrm{vec}^H(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q)\mathbf{e}_{qk}\} & \\ -\mathrm{Tr}(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H) + I_{qk}^{\rm ave} &\geq 0, \\ \forall \mathbf{e}_{qk} : -\mathbf{e}_{qk}^H(\mathbf{T}_{qk}^T \otimes \mathbf{I})\mathbf{e}_{qk} + \varepsilon_{qk}^2 &\geq 0. \end{aligned} \quad (28)$$

where $\mathbf{e}_{qk} \triangleq \mathrm{vec}(\mathbf{E}_{qk})$. Invoking Lemma 1 and following similar steps as in the proof of Proposition 1, one can see that (28) holds if and only if there exists $\mu_{qk} \geq 0$ such that

$$\left[\begin{array}{cc} (\mu_{qk}\mathbf{T}_{qk} - \mathbf{Q}_q)^T \otimes \mathbf{I} & -\mathrm{vec}(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q) \\ -\mathrm{vec}^H(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q) & I_{qk}^{\rm ave} - \mu_{qk}\varepsilon_{qk}^2 \\ & -\mathrm{Tr}(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H) \end{array} \right] \succeq \mathbf{0}. \quad (29)$$

The condition $f_1(\hat{\mathbf{x}}) > 0$, in Lemma 1, is readily satisfied since $\varepsilon_{qk}^2 > 0$.

Let us consider now the constraint

$$\lambda_{\max}((\hat{\mathbf{G}}_{qk} + \mathbf{E}_{qk})\mathbf{Q}_q(\hat{\mathbf{G}}_{qk} + \mathbf{E}_{qk})^H) \leq I_{qk}^{\rm peak}, \ \forall \mathbf{E}_{qk} \in \mathcal{E}_{qk} \quad (30)$$

which is equivalent to

$$\begin{aligned} I_{qk}^{\rm peak}\mathbf{I} - \hat{\mathbf{G}}_{qk}\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H - \hat{\mathbf{G}}_{qk}\mathbf{Q}_q\mathbf{E}_{qk}^H & \\ -\mathbf{E}_{qk}\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H - \mathbf{E}_{qk}\mathbf{Q}_q\mathbf{E}_{qk}^H &\succeq \mathbf{0}, \\ \forall \mathbf{E}_{qk} : \frac{1}{\varepsilon_{qk}^2}\mathrm{Tr}\left(\mathbf{E}_{qk}\mathbf{T}_{qk}\mathbf{E}_{qk}^H\right) &\leq 1. \end{aligned} \quad (31)$$

It follows from Lemma 2 that (31) holds if and only if there exists $\omega_{qk} \geq 0$ such that

$$\left[\begin{array}{cc} I_{qk}^{\rm peak}\mathbf{I} - \omega_{qk}\mathbf{I} - \hat{\mathbf{G}}_{qk}\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H & -\hat{\mathbf{G}}_{qk}\mathbf{Q}_q \\ -\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H & \frac{\omega_{qk}}{\varepsilon_{qk}^2}\mathbf{T}_{qk} - \mathbf{Q}_q \end{array} \right] \succeq \mathbf{0}. \quad (32)$$

Then, by defining $\eta_{qk} \triangleq \omega_{qk}/\varepsilon_{qk}^2$, one can easily rewrite (32) as in (26), which completes the proof. ■

Similarly, the equivalent set (26) consists of a group of LMIs, which makes the study of the robust game $\hat{\mathcal{G}}_{\rm mimo}$ more affordable, as well as the computation of the optimal solution to the single-user optimization problem (16).

IV. Nash Equilibria of the Robust Games

Given a strategic noncooperative game, the existence of a NE in pure strategies is not guaranteed; neither is the uniqueness nor the convergence (e.g., of best-response based algorithms) to an equilibrium when one exists (or even is unique). In this section we provide a positive answer to these key issues for the robust game $\hat{\mathcal{G}}_{\rm siso}$ and $\hat{\mathcal{G}}_{\rm mimo}$ introduced in Sections II-A and II-B, respectively.

The main difficulty in the analysis of the proposed games is that one cannot compute the best-response mapping of each user in closed form in either the SISO case or the MIMO case. Hence, none of the results in [16], [17], [19], [20], [23] based on the machinery of the fixed-point theory can be successfully applied to studying our games. We overcome this issue by exploring the more advanced theory of finite dimensional variational inequalities (VIs). A brief introduction to the VI theory is given in Appendix A, while more details on VI and its application to games can be found in [24], [37], and [38], respectively. The basic rule of using the VI theory is to find a correspondence between the formulated game and a VI problem, and then exploit the well established results on VIs.

A. SISO Frequency-Selective Channels

Before studying the game, we introduce the following notations and definitions. Define $\mathbf{B} \in \mathbb{R}^{Q\times Q}$ as

$$[\mathbf{B}]_{qr} \triangleq \left\{ \begin{array}{ll} 0, & \text{if } q = r \\ \max_n \left\{ \dfrac{|h_{rq}(n)|^2}{|h_{rr}(n)|^2} \cdot \mathrm{innr}_{rq}(n) \right\}, & \text{if } q \neq r \end{array} \right. \quad (33)$$

where the interference-plus-noise to noise ratio $\mathrm{innr}_{rq}(n)$ is defined as

$$\mathrm{innr}_{rq}(n) \triangleq \frac{\sigma_r^2(n) + \sum_{r'=1}^Q |h_{r'r}(n)|^2 \hat{p}_{r'}^{\max}(n)}{\sigma_q^2(n)} \quad (34)$$

with

$$\hat{p}_q^{\max}(n) \triangleq \min\{P_q^{\rm ave}, p_q^{\rm peak}(n), \hat{I}_q^{\rm ave}(n), \hat{I}_q^{\rm peak}(n)\} \quad (35)$$

$$\hat{I}_q^{\rm ave}(n) \triangleq \min_k I_{qk}^{\rm ave} \cdot \min\left\{ \frac{w_{qk}(n)}{\varepsilon_{qk}^2}, \frac{1}{|\hat{g}_{qk}(n)|^2} \right\} \quad (36)$$

$$\hat{I}_q^{\rm peak}(n) \triangleq \min_k \frac{I_{qk}^{\rm peak}(n)}{\left(|\hat{g}_{qk}(n)| + \varepsilon_{qk}/\sqrt{w_{qk}(n)}\right)^2}. \quad (37)$$

The matrix $\mathbf{B}$ in (33) is instrumental to obtain sufficient conditions guaranteeing the uniqueness of the NE as well as the convergence of the proposed algorithms. Following similar steps as in [24], one can rewrite the game $\hat{\mathcal{G}}_{\rm siso}$ as a proper VI problem. Then, building on the equivalent form of $\hat{\mathcal{P}}_q^{\rm int}$ in Proposition 1, we characterize the robust game $\hat{\mathcal{G}}_{\rm siso}$ as given next (The proof is similar to that of [24, Theorem 6] and is omitted because of the space limitation).

Theorem 1: Suppose w.l.o.g. that each strategy set $\hat{\mathcal{P}}_q$ is nonempty. Then, the game $\hat{\mathcal{G}}_{\rm siso}$ admits a NE. Furthermore, the NE of $\hat{\mathcal{G}}_{\rm siso}$ is unique if $\rho(\mathbf{B}) < 1$, where $\mathbf{B}$ is defined in (33). ■

Remark 1: On the uniqueness condition. The uniqueness condition $\rho(\mathbf{B}) < 1$ has an intuitive physical interpretation: The NE is unique if there is not too much interference in the system [15]–[18], [24]. A sufficient condition for $\rho(\mathbf{B}) < 1$ is given by one of the two following conditions:
i) low received interference

$$\sum_{r\neq q} \max_n \left\{ \frac{|h_{rq}(n)|^2}{|h_{rr}(n)|^2} \cdot \mathrm{innr}_{rq}(n) \right\} < 1, \ \forall q \in \Omega \qquad (38)$$

ii) low generated interference

$$\sum_{q\neq r} \max_n \left\{ \frac{|h_{rq}(n)|^2}{|h_{rr}(n)|^2} \cdot \mathrm{innr}_{rq}(n) \right\} < 1, \ \forall r \in \Omega \qquad (39)$$

where the first condition imposes a limit on the interference that each receiver can tolerate, whereas the second one introduces an upper bound on the interference that each transmitter can generate.

Distributed Algorithms: Asynchronous SISO RICA. We focus on distributed algorithms, named robust iterative CR algorithms (RICAs), to reach the NE of $\hat{\mathcal{G}}_{\text{siso}}$. We propose totally asynchronous algorithms, in the sense that in the updating procedure some users may change their strategies more frequently than the others, and they may even use *outdated* information on the interference caused by the others. To formally describe the asynchronous algorithms, we first introduce some definitions as given in [17]. Assume w.l.o.g. that the set of times at which the SUs update their strategies is a discrete set $\mathcal{T} = \mathbb{N}_+ = \{0, 1, 2, \ldots\}$. Let $\mathbf{p}_q^{(m)}$ denote the power allocation of SU q at the m^{th} iteration, and $\mathcal{T}_q \subseteq \mathcal{T}$ denote the set of times m at which $\mathbf{p}_q^{(m)}$ is updated (thus, at time $m \notin \mathcal{T}_q$, $\mathbf{p}^{(m)}$ is left unchanged). Let $\tau_r^q(m)$ denote the most recent time at which the interference from SU r is perceived by SU q at the m^{th} iteration (thus $0 \leq \tau_r^q(m) \leq m$). At the m^{th} iteration, SU q chooses his best response using the interference level caused by $\mathbf{p}_{-q}^{(\tau^q(m))} \triangleq \left(\mathbf{p}_1^{(\tau_1^q(m))}, \ldots, \mathbf{p}_{q-1}^{(\tau_{q-1}^q(m))}, \mathbf{p}_{q+1}^{(\tau_{q+1}^q(m))}, \ldots, \mathbf{p}_Q^{(\tau_Q^q(m))}\right)$. Some standard conditions in asynchronous convergence theory that are fulfilled in any practical implementation need to be satisfied by the schedule $\{\tau_r^q(m)\}$ and $\{\mathcal{T}_q\}$; we refer to [17] for the details. Through the whole paper we assume that these conditions are satisfied and call such an updating schedule as feasible. Using the above notations, the asynchronous SISO RICA is formally described in Algorithm 1. The convergence properties of Algorithm 1 are given in Theorem 1, whose proof follows similar steps of that for [24, Theorem 9] and thus is omitted because of the space limitation.

Algorithm 1 : Asynchronous SISO RICA

1: Choose any feasible $\mathbf{p}_q^{(0)}$, $\forall q \in \Omega$;
2: For $m = 0 : N_{it}$
3: For each $q \in \Omega$, compute:

$$\mathbf{p}_q^{(m+1)} = \begin{cases} \mathbf{p}_q^\star = \arg \max\limits_{\mathbf{p}_q \in \hat{\mathcal{P}}_q} r_q\left(\mathbf{p}_q, \mathbf{p}_{-q}^{(\tau^q(m))}\right), & \text{if } m \in \mathcal{T}_q \\ \mathbf{p}_q^{(m)}, & \text{otherwise} \end{cases}$$

4: End

Theorem 2: Suppose that $\rho(\mathbf{B}) < 1$. Then, any sequence $\{\mathbf{p}^{(m+1)}\}_{m=0}^{\infty}$ generated by the asynchronous SISO RICA described in Algorithm 1 converges to the unique NE of the game $\hat{\mathcal{G}}_{\text{siso}}$ for any feasible updating schedule of the users. ■

Note that Algorithm 1 contains as special cases plenty of algorithms obtained by choosing different schedules of users in the updating procedure. Two special cases are the sequential and simultaneous RICAs, as the counterparts of the well-known sequential and simultaneous iterative water-filling algorithm (IWFA) [13], [16]. However, RICAs differ from the classical IWFAs in the following points: i) RICAs preserve the QoS of PUs, even with imperfect PU CSI; and ii) The single-user optimization problem (8) does not admit a closed-form solution like water-filling. In the section V-A, we will show how to efficiently solve (8).

From Theorems 1 and 2, we can observe an interesting phenomenon: The more uncertainty on PU CSI, the more likely the proposed algorithms converge to the unique NE of the game $\hat{\mathcal{G}}_{\text{siso}}$. To see this, denote by $\mathbf{B}(\boldsymbol{\varepsilon}^{(1)})$ and $\mathbf{B}(\boldsymbol{\varepsilon}^{(2)})$ the matrices defined in (33) for $\boldsymbol{\varepsilon}^{(1)} \leq \boldsymbol{\varepsilon}^{(2)}$, respectively. Then, we have $\mathbf{0} \leq \mathbf{B}(\boldsymbol{\varepsilon}^{(2)}) \leq \mathbf{B}(\boldsymbol{\varepsilon}^{(1)})$ componentwise, implying that $\rho(\mathbf{B}(\boldsymbol{\varepsilon}^{(2)})) \leq \rho(\mathbf{B}(\boldsymbol{\varepsilon}^{(1)}))$. The reason is that, when the uncertainty region of one PU's channel enlarges, the worst-case robust transmit strategy of one SU will restrict more conservatively his interference to this PU, which meanwhile also causes less interference to the others SUs. Observe also that, as $\varepsilon_{qk} \to 0$, $\forall q, k$, the uniqueness and convergence conditions for the robust game $\hat{\mathcal{G}}_{\text{siso}}$ coincide with those for the non-robust game $\mathcal{G}_{\text{siso}}$ as in [24].

B. MIMO Channels

Differently from the game $\hat{\mathcal{G}}_{\text{siso}}$ over SISO frequency-selective channels, in the game $\hat{\mathcal{G}}_{\text{mimo}}$, the strategies of each player are complex matrices, making $\hat{\mathcal{G}}_{\text{mimo}}$ much more difficult to analyze. To simplify the study of the game, we first transform $\hat{\mathcal{G}}_{\text{mimo}}$ into an equivalent game with real strategy sets by the isomorphism mapping between complex and real matrices. More specifically, the real counterparts, under the isomorphism mapping, of a complex matrix $\mathbf{X} = \mathbf{X}_R + i\mathbf{X}_I$ and a complex vector $\mathbf{x} = \mathbf{x}_R + i\mathbf{x}_I$ are given by

$$\breve{\mathbf{X}} = \begin{bmatrix} \mathbf{X}_R & -\mathbf{X}_I \\ \mathbf{X}_I & \mathbf{X}_R \end{bmatrix} \text{ and } \breve{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix} \qquad (40)$$

respectively. A variety of properties of the isomorphism mapping can be found in [45].

Now, consider the game $\breve{\mathcal{G}}_{\text{mimo}} = \langle \Omega, \{\breve{\mathcal{Q}}_q\}_{q\in\Omega}, \{\breve{R}_q\}_{q\in\Omega} \rangle$, where each player's strategy is given by the real matrix

$$\breve{\mathbf{Q}}_q \triangleq \begin{bmatrix} \breve{\mathbf{Q}}_{q,11} & \breve{\mathbf{Q}}_{q,12} \\ \breve{\mathbf{Q}}_{q,21} & \breve{\mathbf{Q}}_{q,22} \end{bmatrix} \in \mathbb{R}^{2N\times 2N}$$

with $\breve{\mathbf{Q}}_{q,ij} \in \mathbb{R}^{N\times N}$, $i, j = 1, 2$. Denote by $\breve{\mathbf{H}}_{rq}$, $\breve{\mathbf{G}}_{qk}$, $\breve{\mathbf{T}}_{qk}$, $\breve{\mathbf{R}}_{n_q}$ the real counterparts of the complex matrices $\mathbf{H}_{rq}$, $\hat{\mathbf{G}}_{qk}$, $\mathbf{T}_{qk}$, $\mathbf{R}_{n_q}$, respectively. For each player q, the utility function $\breve{R}_q$ is defined as

$$\breve{R}_q(\breve{\mathbf{Q}}_q, \breve{\mathbf{Q}}_{-q}) \triangleq \log\det(\breve{\mathbf{R}}_{-q}(\breve{\mathbf{Q}}_{-q}) + \breve{\mathbf{H}}_{qq}\breve{\mathbf{Q}}_q\breve{\mathbf{H}}_{qq}^T) \qquad (41)$$

where $\breve{\mathbf{R}}_{-q}(\breve{\mathbf{Q}}_{-q}) \triangleq \breve{\mathbf{R}}_{n_q} + \sum_{r\neq q} \breve{\mathbf{H}}_{rq}\breve{\mathbf{Q}}_r\breve{\mathbf{H}}_{rq}^T$. The strategy set $\breve{\mathcal{Q}}_q$ is given by (42) at the bottom of the page, where

$$\mathbf{P} \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}.$$

In the game $\breve{\mathcal{G}}_{\rm mimo}$, the optimal strategy of each SU q, given the strategies of the others, is the solution to the following convex optimization problem:

$$\underset{\breve{\mathbf{Q}}_q \in \breve{\mathcal{Q}}_q}{\text{maximize}} \ \breve{R}_q(\breve{\mathbf{Q}}_q, \breve{\mathbf{Q}}_{-q}). \tag{43}$$

The relationship between $\breve{\mathcal{G}}_{\rm mimo}$ and $\hat{\mathcal{G}}_{\rm mimo}$ is given next.

Lemma 3: The games $\breve{\mathcal{G}}_{\rm mimo}$ and $\hat{\mathcal{G}}_{\rm mimo}$ are equivalent, and the equivalence is in the following sense. Let $\breve{\mathbf{Q}}^\star \triangleq (\breve{\mathbf{Q}}_q^\star)_{q=1}^Q$ be a NE of $\breve{\mathcal{G}}_{\rm mimo}$. Then, $\mathbf{Q}^\star$ is a NE of $\hat{\mathcal{G}}_{\rm mimo}$ if and only if $\breve{\mathbf{Q}}^\star$ is a NE of $\breve{\mathcal{G}}_{\rm mimo}$ under the one-to-one mapping ${\rm Re}(\mathbf{Q}_q) = \breve{\mathbf{Q}}_{q,11}$ and ${\rm Im}(\mathbf{Q}_q) = \breve{\mathbf{Q}}_{q,21}$ for $q = 1, \ldots, Q$.

Proof: The equivalence follows readily from the properties of the isomorphism mapping [45]. To be specific, let $\breve{\mathbf{X}}$, $\breve{\mathbf{A}}$ and $\breve{\mathbf{B}}$ be the real counterparts of complex matrices $\mathbf{X}$, $\mathbf{A}$ and $\mathbf{B}$, respectively. Then, 1) $\mathbf{X} = \mathbf{A} + \mathbf{B} \Leftrightarrow \breve{\mathbf{X}} = \breve{\mathbf{A}} + \breve{\mathbf{B}}$; 2) $\mathbf{X} = \mathbf{A}\mathbf{B} \Leftrightarrow \breve{\mathbf{X}} = \breve{\mathbf{A}}\breve{\mathbf{B}}$; 3) $\mathbf{X} = \mathbf{B}\mathbf{A}\mathbf{B}^H \Leftrightarrow \breve{\mathbf{X}} = \breve{\mathbf{B}}\breve{\mathbf{A}}\breve{\mathbf{B}}^T$; 4) $\mathbf{X} = \mathbf{A}\otimes\mathbf{I} \Leftrightarrow \breve{\mathbf{X}} = \breve{\mathbf{A}}\otimes\mathbf{I}$; 5) $\mathbf{X} = \mathbf{A}^T \Leftrightarrow \breve{\mathbf{X}} = \mathbf{P}\breve{\mathbf{A}}^T\mathbf{P}$, where $\mathbf{P}$ is defined in (IV-B). Suppose that $\mathbf{X}$ and $\mathbf{A}$ are square matrices. Then, 6) $\mathbf{X} = \mathbf{A}^{-1} \Leftrightarrow \breve{\mathbf{X}} = \breve{\mathbf{A}}^{-1}$; 7) ${\rm Tr}(\breve{\mathbf{X}}) = 2{\rm Tr}(\mathbf{X})$; 8) $\det(\breve{\mathbf{X}}) = |\det(\mathbf{X})|^2$; 9) $\breve{\mathbf{X}} \succeq \mathbf{0} \Leftrightarrow \mathbf{X} \succeq \mathbf{0}$; 10)

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{X} \end{bmatrix} \succeq \mathbf{0} \Leftrightarrow \begin{bmatrix} \breve{\mathbf{A}} & \breve{\mathbf{B}} \\ \breve{\mathbf{B}}^T & \breve{\mathbf{X}} \end{bmatrix} \succeq \mathbf{0};$$

and 11) λ is an eigenvalue of $\breve{\mathbf{X}}$ if and only if it is an eigenvalue of $\mathbf{X}$. ■

Lemma 3 implies that we can study w.l.o.g. $\breve{\mathcal{G}}_{\rm mimo}$ instead of $\hat{\mathcal{G}}_{\rm mimo}$. To this end, we first introduce some definitions. Let $\mathbf{D}_\alpha, \mathbf{B}_\beta \in \mathbb{R}^{Q\times Q}$ be defined as

$$\mathbf{D}_\alpha \triangleq {\rm diag}\left\{\alpha_q^{\min}\right\}_{q=1}^Q, \ [\mathbf{B}_\beta]_{qr} \triangleq \begin{cases} 0, & \text{if } r = q \\ \beta_{rq}^{\max}, & \text{if } r \neq q \end{cases} \tag{44}$$

where

$$\alpha_q^{\min} \triangleq \lambda_{\min}^2\left(\mathbf{H}_{qq}^H\left(\mathbf{R}_{n_q} + \sum_{r=1}^Q \gamma_r \mathbf{H}_{rq}\mathbf{H}_{rq}^H\right)^{-1}\mathbf{H}_{qq}\right) \tag{45}$$

$$\beta_{rq}^{\max} \triangleq \lambda_{\max}\left(\mathbf{H}_{rq}\mathbf{H}_{rq}^H\right)\lambda_{\max}\left(\mathbf{H}_{qq}^H\mathbf{R}_{n_q}^{-2}\mathbf{H}_{qq}\right) \tag{46}$$

with

$$\gamma_q \triangleq \min\left\{P_q^{\rm ave}, P_q^{\rm peak}, \hat{I}_q\right\} \tag{47}$$

$$\hat{I}_q \triangleq \min_k \min\left\{I_{qk}^{\rm ave}, I_{qk}^{\rm peak}\right\} \cdot \min\left\{\frac{\lambda_{\max}(\mathbf{T}_{qk})}{\varepsilon_{qk}^2}, \frac{1}{\lambda_{\min}(\hat{\mathbf{G}}_{qk}\hat{\mathbf{G}}_{qk}^H)}\right\} \tag{48}$$

The main properties of game $\breve{\mathcal{G}}_{\rm mimo}$ and thus $\hat{\mathcal{G}}_{\rm mimo}$ are given in the following.

Theorem 3: Given the game $\hat{\mathcal{G}}_{\rm mimo}$ (or $\breve{\mathcal{G}}_{\rm mimo}$), the following holds: a) Suppose w.l.o.g. that each strategy set $\hat{\mathcal{Q}}_q$ (or $\breve{\mathcal{Q}}_q$) is nonempty. Then, the game admits a NE. b) Suppose that each channel matrix $\mathbf{H}_{qq}$ is full column-rank. Then, the NE of the game is unique if $\rho(\mathbf{D}_\alpha^{-1}\mathbf{B}_\beta) < 1$, where $\mathbf{D}_\alpha$ and $\mathbf{B}_\beta$ are defined in (44).

Proof: See Appendix A. ■

Distributed algorithms: Asynchronous MIMO RICA. We propose totally asynchronous and distributed algorithms to reach the NE of $\hat{\mathcal{G}}_{\rm mimo}$. Similar to the SISO case, let us introduce the set $\mathcal{T}_q$ and the time index $\tau_r^q(m)$, and define $\mathbf{Q}_{-q}^{(\tau^q(m))} \triangleq \left(\mathbf{Q}_1^{(\tau_1^q(m))}, \ldots, \mathbf{Q}_{q-1}^{(\tau_{q-1}^q(m))}, \mathbf{Q}_{q+1}^{(\tau_{q+1}^q(m))}, \ldots, \mathbf{Q}_Q^{(\tau_Q^q(m))}\right)$. Then, the asynchronous MIMO RICA is formally described in Algorithm 2, whose convergence properties are given in Theorem 4.

Algorithm 2 : Asynchronous MIMO RICA

1: Choose any feasible $\mathbf{Q}_q^{(0)}$, $\forall q \in \Omega$;
2: For $m = 0 : N_{it}$
3: For each $q \in \Omega$, compute:
$\mathbf{Q}_q^{(m+1)} =$
$$\begin{cases} \mathbf{Q}_q^\star \in \arg\max\limits_{\mathbf{Q}_q \in \hat{\mathcal{Q}}_q} R_q\left(\mathbf{Q}_q, \mathbf{Q}_{-q}^{(\tau^q(m))}\right), & \text{if } m \in \mathcal{T}_q \\ \mathbf{Q}_q^{(m)}, & \text{otherwise} \end{cases}$$
4: End

Theorem 4: Suppose that each channel matrix $\mathbf{H}_{qq}$ is full column-rank, and $\rho(\mathbf{D}_\alpha^{-1}\mathbf{B}_\beta) < 1$. Then, any sequence $\{\mathbf{Q}^{(m+1)}\}_{m=0}^\infty$ generated by the asynchronous MIMO RICA described in Algorithm 2 converges to the unique NE of $\hat{\mathcal{G}}_{\rm mimo}$ for any feasible updating schedule of the users.

Proof: See Appendix B. ■

Note that the condition $\rho(\mathbf{D}_\alpha^{-1}\mathbf{B}_\beta) < 1$ is satisfied, similar to the SISO case, when there is not too much interference

$$\breve{\mathcal{Q}}_q \triangleq \left\{ \begin{array}{ll} \breve{\mathbf{Q}}_q \in \mathbb{S}_+^{2N}: & \exists \boldsymbol{\mu}_q \triangleq (\mu_{qk})_{k=1}^K \geq \mathbf{0}, \ \exists \boldsymbol{\eta}_q \triangleq (\eta_{qk})_{k=1}^K \geq \mathbf{0}, \\ & \begin{bmatrix} \mathbf{P}(\mu_{qk}\breve{\mathbf{T}}_{qk} - \breve{\mathbf{Q}}_q)^T\mathbf{P}\otimes\mathbf{I} & -{\rm vec}(\breve{\mathbf{G}}_{qk}\mathbf{P}\breve{\mathbf{Q}}_q\mathbf{P}) \\ -{\rm vec}^T(\breve{\mathbf{G}}_{qk}\mathbf{P}\breve{\mathbf{Q}}_q\mathbf{P}) & I_{qk}^{\rm ave} - \mu_{qk}\varepsilon_{q,k}^2 - \frac{1}{2}{\rm Tr}(\breve{\mathbf{G}}_{qk}\breve{\mathbf{Q}}_q\breve{\mathbf{G}}_{qk}^T) \end{bmatrix} \succeq \mathbf{0}, \\ & \begin{bmatrix} I_{qk}^{\rm peak}\mathbf{I} - \eta_{qk}\varepsilon_{gk}^2\mathbf{I} - \breve{\mathbf{G}}_{qk}\breve{\mathbf{Q}}_q\breve{\mathbf{G}}_{qk}^T & -\breve{\mathbf{G}}_{qk}\breve{\mathbf{Q}}_q \\ -\breve{\mathbf{Q}}_q\breve{\mathbf{G}}_{qk}^T & \eta_{qk}\breve{\mathbf{T}}_{qk} - \breve{\mathbf{Q}}_q \end{bmatrix} \succeq \mathbf{0}, \ \forall k, \\ & {\rm Tr}(\breve{\mathbf{Q}}_q) \leq 2P_q^{\rm ave}, \ \lambda_{\max}(\breve{\mathbf{Q}}_q) \leq P_q^{\rm peak}, \\ & \breve{\mathbf{Q}}_q = \breve{\mathbf{Q}}_q^T, \ \breve{\mathbf{Q}}_{q,11} = \breve{\mathbf{Q}}_{q,22}, \ \breve{\mathbf{Q}}_{q,21} = -\breve{\mathbf{Q}}_{q,12} \end{array} \right\} \tag{42}$$

among the SUs [19], [20], [23]. One can also conclude from Theorems 3 and 4 that the worst-case robust design turns channel uncertainty to a conducive factor for the SU network to converge to the unique NE of $\hat{\mathcal{G}}_{\text{mimo}}$. The results in Theorems 3 and 4 can be readily applied to the non-robust game $\mathcal{G}_{\text{mimo}}$ by taking $\varepsilon_{qk} \to 0$, $\forall q, k$.

V. Best Response of Single-User Optimization

In the distributed algorithms described so far, each SU needs to find his best response by solving the single-user optimization problem (8) in the game $\hat{\mathcal{G}}_{\text{siso}}$, or (16) in the game $\hat{\mathcal{G}}_{\text{mimo}}$. The problems (8) and (16), even though convex, are still difficult to handle, due to the intractable forms of $\hat{\mathcal{P}}_q^{\text{int}}$ and $\hat{\mathcal{Q}}_q^{\text{int}}$ defined in (6) and (14), respectively. Nevertheless, this difficulty can be overcome by using the equivalent forms of $\hat{\mathcal{P}}_q^{\text{int}}$ and $\hat{\mathcal{Q}}_q^{\text{int}}$ in (17) and (26), respectively. In this section, we provide efficient methods to solve (8) and (16).

A. SISO Frequency-Selective Channels

According to Proposition 1, the problem (8) is equivalent to (49) at the bottom of the page with the auxiliary variable $\boldsymbol{\mu}_q$. This is a tractable convex problem, since its feasible set is given by an intersection of the sublevels of the linear functions. Consequently, many general numerical methods, e.g., the interior-point method [43], can be applied to finding the optimal solution to (49) in polynomial time. Interestingly, one can simplify the computation of the optimal solution of (49) by exploring the structure of the problem and rewriting it in a more convenient form, as detailed next.

Proposition 3: The problem (49) is equivalent to (50) at the bottom of the page.

Proof: We will use the following fact.

Lemma 4 (Schur's complement [39]): Let

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B}^H \\ \mathbf{B} & \mathbf{C} \end{bmatrix}$$

be a Hermitian matrix. Then, $\mathbf{M} \succeq \mathbf{0}$ if and only if $\mathbf{A} - \mathbf{B}^H\mathbf{C}^{-1}\mathbf{B} \succeq \mathbf{0}$ (assuming $\mathbf{C} \succ \mathbf{0}$), or $\mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^H \succeq \mathbf{0}$ (assuming $\mathbf{A} \succ \mathbf{0}$). □

Let $\mathbf{P}_q \triangleq \text{diag}\{\mathbf{p}_q\}$ and $\mathbf{W}_{qk} \triangleq \text{diag}\{\mathbf{w}_{qk}\}$. We first notice that the LMI in (49) implies $\mu_{qk}\mathbf{W}_{qk} - \mathbf{P}_q \succeq \mathbf{0}$, i.e., $\mu_{qk}w_{qk}(n) \geq p_q(n)$, $\forall n$. We consider two cases: i) $\mu_{qk}w_{qk}(n) > p_q(n)$ for all n; ii) $\mu_{qk}w_{qk}(m) = p_q(m)$ for some m. In case i), $\mu_{qk}\mathbf{W}_{qk} - \mathbf{P}_q$ is invertible, so the LMI in (49), from Lemma 4, amounts to

$$\begin{aligned} & I_{qk}^{\text{ave}} - \mu_{qk}\varepsilon_{qk}^2 - \hat{\boldsymbol{g}}_{qk}^H\mathbf{P}_q\hat{\boldsymbol{g}}_{qk} \\ & \qquad - \hat{\boldsymbol{g}}_{qk}^H\mathbf{P}_q(\mu_{qk}\mathbf{W}_{qk} - \mathbf{P}_q)^{-1}\mathbf{P}_q\hat{\boldsymbol{g}}_{qk} \\ = \; & I_{qk}^{\text{ave}} - \mu_{qk}\varepsilon_{qk}^2 - \sum_{n=1}^{N} |\hat{g}_{qk}(n)|^2 p_q(n) \\ & \qquad - \sum_{n=1}^{N} |\hat{g}_{qk}(n)|^2 \frac{p_q^2(n)}{\mu_{qk}w_{qk}(n) - p_q(n)} \\ = \; & I_{qk}^{\text{ave}} - \mu_{qk}\varepsilon_{qk}^2 - \sum_{n=1}^{N} \frac{\mu_{qk}w_{qk}(n)|\hat{g}_{qk}(n)|^2 p_q(n)}{\mu_{qk}w_{qk}(n) - p_q(n)} \\ \geq \; & 0. \end{aligned} \tag{51}$$

In case ii), by selecting the mth and $(N+1)$th rows and columns of the LMI in (49), we have

$$\begin{bmatrix} 0 & -\hat{g}_{qk}(m)p_q(m) \\ -\hat{g}_{qk}^*(m)p_q(m) & I_{qk}^{\text{ave}} - \mu_{qk}\varepsilon_{qk}^2 - \hat{\boldsymbol{g}}_{qk}^H\mathbf{P}_q\hat{\boldsymbol{g}}_{qk} \end{bmatrix} \succeq \mathbf{0} \tag{52}$$

implying $p_q(m) = \mu_{qk} = 0$. However, if $\mu_{qk} = 0$, then $p_q(n) = 0$ for $\forall n$, meaning that at the optimum of (49) it must be $\mu_{qk}w_{qk}(n) > p_q(n)$ for all n. In other words, case b) cannot happen in practice. Therefore, the equivalence between (51) and the LMI in (49) can be extended from the case that $\mu_{qk}w_{qk}(n) > p_q(n)$ for $\forall n$ to the case that $\mu_{qk}w_{qk}(n) \geq p_q(n)$ for $\forall n$, without losing any optimality. This completes the proof. ■

The equivalent reformulation (50) has several desirable properties. First, it is still a convex problem. Indeed, the function

$$\varphi_{nqk}(p_q(n), \mu_{qk}) \triangleq \frac{\mu_{qk}w_{qk}(n)|\hat{g}_{qk}(n)|^2 p_q(n)}{\mu_{qk}w_{qk}(n) - p_q(n)} \tag{53}$$

$$\begin{aligned} \underset{\mathbf{p}_q, \boldsymbol{\mu}_q}{\text{maximize}} \quad & r_q(\mathbf{p}_q, \mathbf{p}_{-q}) \\ \text{subject to} \quad & \mathbf{p}_q \in \mathcal{P}_q^{\text{pow}} \\ & \begin{bmatrix} \text{diag}\{\mu_{qk}\mathbf{w}_{qk} - \mathbf{p}_q\} & -\mathbf{p}_q \odot \hat{\boldsymbol{g}}_{qk} \\ -(\mathbf{p}_q \odot \hat{\boldsymbol{g}}_{qk})^H & I_{qk}^{\text{ave}} - \mu_{qk}\varepsilon_{qk}^2 - \hat{\boldsymbol{g}}_{qk}^H\text{diag}\{\mathbf{p}_q\}\hat{\boldsymbol{g}}_{qk} \end{bmatrix} \succeq \mathbf{0}, \ \forall k \\ & \left(|\hat{g}_{qk}(n)| + \varepsilon_{qk}/\sqrt{w_{qk}(n)}\right)^2 p_q(n) \leq I_{qk}^{\text{peak}}(n), \ \forall n, k \end{aligned} \tag{49}$$

$$\begin{aligned} \underset{\mathbf{p}_q, \boldsymbol{\mu}_q}{\text{maximize}} \quad & r_q(\mathbf{p}_q, \mathbf{p}_{-q}) \\ \text{subject to} \quad & \mathbf{p}_q \in \mathcal{P}_q^{\text{pow}} \\ & \textstyle\sum_{n=1}^{N} \frac{\mu_{qk}w_{qk}(n)|\hat{g}_{qk}(n)|^2 p_q(n)}{\mu_{qk}w_{qk}(n) - p_q(n)} + \mu_{qk}\varepsilon_{qk}^2 \leq I_{qk}^{\text{ave}}, \ \forall k \\ & p_q(n) \leq \mu_{qk}w_{qk}(n), \ \forall n, k \\ & \left(|\hat{g}_{qk}(n)| + \varepsilon_{qk}/\sqrt{w_{qk}(n)}\right)^2 p_q(n) \leq I_{qk}^{\text{peak}}(n), \ \forall n, k. \end{aligned} \tag{50}$$

is jointly convex in $(p_q(n), \mu_{qk})$ (as its Hessian matrix is positive semidefinite), implying that the constraint

$$\sum_{n=1}^{N} \varphi_{nqk}(p_q(n), \mu_{qk}) + \mu_{qk}\varepsilon_{qk}^2 \le I_{qk}^{\text{ave}} \tag{54}$$

is convex. Second, compared to (49), the problem (50) needs less computational complexity, since the dimension of the LMI constraints in (49) is much bigger than that of the equivalent constraints in (50). Third, (50) has a hidden decomposable structure in terms of the variables $\{p(n)\}_{n=1}^{N}$. Therefore, it is possible, using a multilevel decomposition method [8], e.g., the dual-primal decomposition, to decompose (50) to several levels, each containing a simpler problem with the dimension much smaller than that of (50). The details on how to use the decomposition method to solve (50) are provided in Appendix C.

B. MIMO Channels

Invoking Proposition 2, the problem (16) can be equivalently rewritten as (55) at the bottom of the page with the auxiliary variables $\boldsymbol{\mu}_q$ and $\boldsymbol{\eta}_q$. Similarly, (55) is a convex problem with LMI constraints, and thus can be efficiently solved. Unlike (49) in the SISO case, in general (55) cannot be simplified, in order to reduce the computational complexity. Nevertheless, in the special case of MISO channels–the receivers of SUs and PUs are equipped with only one antenna–the original problem (16) can be equivalently transformed into a second-order cone program (SOCP), thus leading to lower complexity.

Denote the MISO channel from SU r' transmitter to SU q' receiver by $\mathbf{h}_{rq} \triangleq \mathbf{H}_{rq}^H \in \mathbb{C}^{N_r \times 1}$, and the nominal MISO channel from SU q' transmitter to PU k by $\hat{\mathbf{g}}_{qk} \triangleq \hat{\mathbf{G}}_{qk}^H \in \mathbb{C}^{N_q \times 1}$. Let $P_q \triangleq \min\{P_q^{\text{ave}}, P_q^{\text{peak}}\}$ and $I_{qk} \triangleq \min\{I_{qk}^{\text{ave}}, I_{qk}^{\text{peak}}\}$. Then, we have the following result.

Proposition 4: In the MISO case, the optimal solution to (16) is given by $\mathbf{Q}_q^\star = \mathbf{q}_q^\star \mathbf{q}_q^{\star H}$, where $\mathbf{q}_q^\star$ is the optimal solution to the following problem:

$$\begin{array}{ll} \underset{\mathbf{q}_q}{\text{maximize}} & \text{Re}(\mathbf{h}_{qq}^H \mathbf{q}_q) \\ \text{subject to} & \text{Im}(\mathbf{h}_{qq}^H \mathbf{q}_q) = 0,\ \|\mathbf{q}_q\| \le \sqrt{P_q} \\ & \left|\hat{\mathbf{g}}_{qk}^H \mathbf{q}_q\right| + \left\|\mathbf{T}_{qk}^{-1/2}\mathbf{q}_q\right\| \le \sqrt{I_{qk}},\ \forall k. \end{array} \tag{56}$$

Proof: In the MISO case, the problem (16) with the robust interference constraints reduces to

$$\begin{array}{ll} \underset{\mathbf{Q}_q \succeq 0}{\text{maximize}} & \mathbf{h}_{qq}^H \mathbf{Q} \mathbf{h}_{qq} \\ \text{subject to} & \text{Tr}(\mathbf{Q}_q) \le P_q^{\text{ave}},\ \lambda_{\max}(\mathbf{Q}_q) \le P_q^{\text{peak}} \\ & (\hat{\mathbf{g}}_{qk} + \mathbf{e}_{qk})^H \mathbf{Q}_q (\hat{\mathbf{g}}_{qk} + \mathbf{e}_{qk}) \le I_{qk}, \\ & \qquad \forall \mathbf{e}_{qk} : \left\|\mathbf{T}_{qk}^{1/2}\mathbf{e}_{qk}\right\|^2 \le \varepsilon_{qk}^2,\ \forall k. \end{array} \tag{57}$$

By interpreting each robust interference constraint as a infinite number of constraints and analyzing the KKT conditions of (57), it can be proved that the solution to (57) is always rank-one (see, e.g., [6] for a similar approach). Then, without loss of any optimality, we can express $\mathbf{Q}_q$ as $\mathbf{Q}_q = \mathbf{q}_q \mathbf{q}_q^H$ and rewrite (57) as

$$\begin{array}{ll} \underset{\|\mathbf{q}_q\|^2 \le P_q}{\text{maximize}} & |\mathbf{h}_{qq}^H \mathbf{q}_q| \\ \text{subject to} & \left|(\hat{\mathbf{g}}_{qk} + \mathbf{e}_{qk})^H \mathbf{q}_q\right|^2 \le I_{qk}, \\ & \qquad \forall \mathbf{e}_{qk} : \left\|\mathbf{T}_{qk}^{1/2}\mathbf{e}_{qk}\right\| \le \varepsilon_{qk},\ \forall k. \end{array} \tag{58}$$

Now we show that (58) can be equivalently reformulated as an SOCP, following similar steps as in [27].

Defining $\hat{\mathbf{e}}_{qk} \triangleq \mathbf{T}_{qk}^{1/2}\mathbf{e}_k$, then the k^{th} robust constraint in (58) is equal to

$$\max_{\|\hat{\mathbf{e}}_{qk}\| \le \varepsilon_{qk}} \left|\hat{\mathbf{g}}_{qk}^H \mathbf{q}_q + \hat{\mathbf{e}}_{qk}^H \mathbf{T}_{qk}^{-1/2} \mathbf{q}_q\right| \le \sqrt{I_{qk}}. \tag{59}$$

Using the triangle inequality and the Cauchy–Schwarz inequality with $\|\hat{\mathbf{e}}_{q,k}\| \le \varepsilon_{q,k}$, it follows that

$$\begin{aligned} & \left|\hat{\mathbf{g}}_{qk}^H \mathbf{q}_q + \hat{\mathbf{e}}_{qk}^H \mathbf{T}_{qk}^{-1/2} \mathbf{q}_q\right| \\ \le\ & \left|\hat{\mathbf{g}}_{qk}^H \mathbf{q}_q\right| + \left|\hat{\mathbf{e}}_{qk}^H \mathbf{T}_{qk}^{-1/2} \mathbf{q}_q\right| \\ \le\ & \left|\hat{\mathbf{g}}_{qk}^H \mathbf{q}_q\right| + \varepsilon_{qk} \left\|\mathbf{T}_{qk}^{-1/2} \mathbf{q}_q\right\| \end{aligned} \tag{60}$$

where the equalities are achieved when

$$\hat{\mathbf{e}}_{qk} = \varepsilon_{qk} e^{j\phi_{qk}} \frac{\mathbf{T}_{qk}^{-1/2}\mathbf{q}_q}{\left\|\mathbf{T}_{qk}^{-1/2}\mathbf{q}_q\right\|} \tag{61}$$

with $\phi_{qk} = \angle(\hat{\mathbf{g}}_{qk}^H \mathbf{q}_q)$. This indicates

$$\max_{\|\hat{\mathbf{e}}_{qk}\| \le \varepsilon_{qk}} \left|\hat{\mathbf{g}}_{qk}^H \mathbf{q}_q + \hat{\mathbf{e}}_{qk}^H \mathbf{T}_{qk}^{-1/2} \mathbf{q}_q\right| = \left|\hat{\mathbf{g}}_{qk}^H \mathbf{q}_q\right| + \varepsilon_{qk} \left\|\mathbf{T}_{qk}^{-1/2} \mathbf{q}_q\right\| \tag{62}$$

so (58) is equivalent to

$$\begin{array}{ll} \underset{\|\mathbf{q}_q\| \le \sqrt{P_q}}{\text{maximize}} & \left|\mathbf{h}_{qq}^H \mathbf{q}_q\right| \\ \text{subject to} & \left|\hat{\mathbf{g}}_{qk}^H \mathbf{q}_q\right| + \varepsilon_{qk} \left\|\mathbf{T}_{qk}^{-1/2} \mathbf{q}_q\right\| \le \sqrt{I_{qk}},\ \forall k. \end{array} \tag{63}$$

$$\begin{array}{ll} \underset{\mathbf{Q}_q, \boldsymbol{\mu}_q, \boldsymbol{\eta}_q}{\text{maximize}} & R_q(\mathbf{Q}_q, \mathbf{Q}_{-q}) \\ \text{subject to} & \mathbf{Q}_q \in \mathcal{Q}_q^{\text{pow}} \\ & \begin{bmatrix} (\mu_{qk}\mathbf{T}_{qk} - \mathbf{Q}_q)^T \otimes \mathbf{I} & -\text{vec}(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q) \\ -\text{vec}^H(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q) & I_{qk}^{\text{ave}} - \mu_{qk}\varepsilon_{qk}^2 - \text{Tr}(\hat{\mathbf{G}}_{qk}\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H) \end{bmatrix} \succeq \mathbf{0} \\ & \begin{bmatrix} I_{qk}^{\text{peak}}\mathbf{I} - \eta_{qk}\varepsilon_{qk}^2\mathbf{I} - \hat{\mathbf{G}}_{qk}\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H & -\hat{\mathbf{G}}_{qk}\mathbf{Q}_q \\ -\mathbf{Q}_q\hat{\mathbf{G}}_{qk}^H & \eta_{qk}\mathbf{T}_{qk} - \mathbf{Q}_q \end{bmatrix} \succeq \mathbf{0},\ \forall k \end{array} \tag{55}$$

Notice that the objective function and constraints in (63) are invariant to the phase rotation on $\mathbf{q}_q$. Therefore, one can choose w.l.o.g. $\mathbf{q}_q$ such that $\mathrm{Re}\{\mathbf{h}_{qq}^H\mathbf{q}_q\} \geq 0$ and $\mathrm{Im}\{\mathbf{h}_{qq}^H\mathbf{q}_q\} = 0$, meaning that (63) is equivalent to (56). ■

Proposition 4 indicates that, in the MISO case, the robust MIMO precoding reduces to the robust beamforming, which can be obtained by solving the SOCP (56). On the other hand, one may notice that the sufficient conditions guaranteeing the convergence of the distributed algorithms to the unique NE of $\hat{\mathcal{G}}_{\text{mimo}}$ are not satisfied in the MISO case, since $\mathbf{H}_{qq}$ is not full column-rank. However, in the MISO case, the best response of each SU–the optimal solution to the problem (57)–is a dominant strategy,[3] implying that the game has a unique NE. Such a NE of $\hat{\mathcal{G}}_{\text{mimo}}$ can be immediately obtained after one round of the decoupled single-user optimization. Finally, we point out that [34] considered, as a special case of our framework, a MISO robust design with only one PU ($K = 1$) and one SU ($Q = 1$), but interestingly provided a semi-closed-form solution.

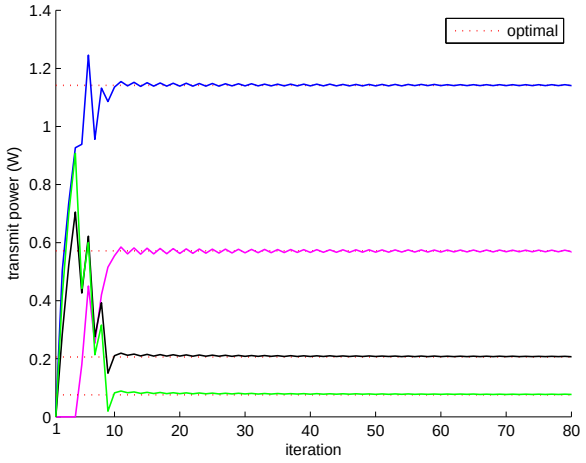


Fig. 2. Iteration of the projection subgradient method in the dual-primal decomposition to find the robust power allocation for a single SU in SISO frequency-selective channels; $N = 4$, $I^{\text{ave}} = -10\text{dB}$, $d = 3$, $\text{SNR} = 3\text{dB}$, $s = 0.04$.

VI. Numerical Results

In this section, we demonstrate the effect of the robust CR design through numerical experiments. For the sake of simplicity, we consider a single PU ($K = 1$) and adopt the following setting for each SU. We consider: 1) average power constraints only and no peak power constraint ($P_q^{\text{peak}}, p_q^{\text{peak}}(n) = \infty$, $\forall q, n$); 2) spherical channel uncertainty regions ($\mathbf{w}_{qk} = \mathbf{1}$, $\mathbf{T}_{qk} = \mathbf{I}$, $\forall q, k$); 3) normalized and uncorrelated noise ($\sigma_q^2(n) = 1$, $\mathbf{R}_{n_q} = \mathbf{I}$, $\forall q, n$); 4) the radii of the uncertainty regions $\mathcal{D}_{qk}$ and $\mathcal{E}_{qk}$ to be a fraction of the nominal channels as $\varepsilon_{qk}^2 = s||\hat{\mathbf{g}}_{qk}||^2$ and $\varepsilon_{qk}^2 = s||\hat{\mathbf{G}}_{qk}||_F^2$ with a common $s \in [0, 1)$, $\forall q, k$. Following the worst-case robustness philosophy, we compare the robust designs ($\hat{\mathcal{G}}_{\text{siso}}$ and $\hat{\mathcal{G}}_{\text{mimo}}$) and the non-robust designs ($\mathcal{G}_{\text{siso}}$ and $\mathcal{G}_{\text{mimo}}$) through their worst-case performance, i.e., the worst-case interference generated at the primary receiver. To this end, we need to know the worst channel error for a specific transmit strategy, which is approximately obtained by choosing the worst one (resulting in the maximum interference) among 1000 randomly generated errors on the boundary[4] of the spherical uncertainty region. In the following, we first show the performance of the robust transmit strategies, obtained in Section V, for a single SU, and then the robust CR systems based on $\hat{\mathcal{G}}_{\text{siso}}$ and $\hat{\mathcal{G}}_{\text{mimo}}$ for multiple SUs.

A. Single-User Scenario

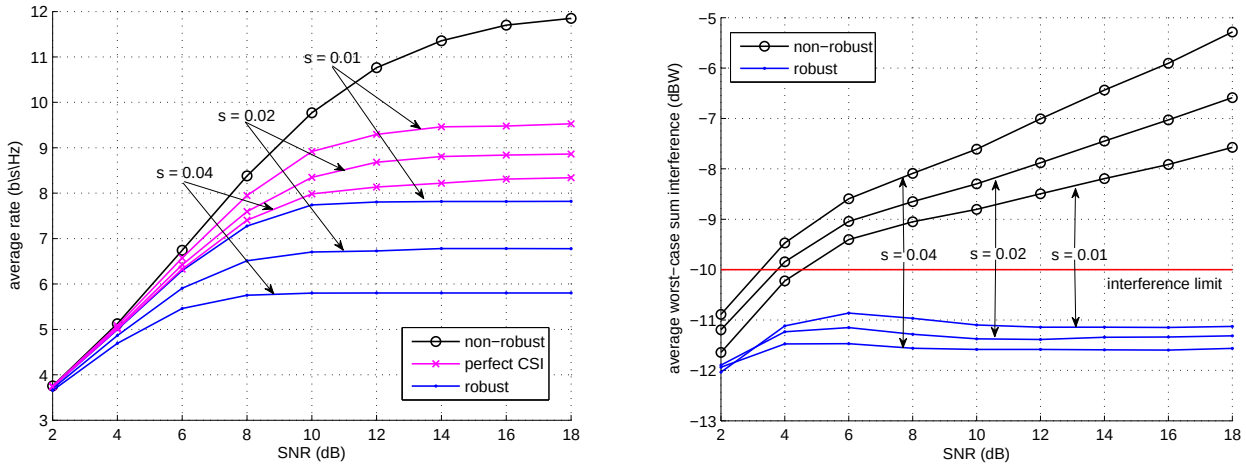
In this scenario, we focus w.l.o.g. on the sum (aggregate) interference generated by the SU, and set $I^{\text{peak}}, I^{\text{peak}}(n) = \infty$, $\forall n$, (the subscripts q and k are suppressed since $Q = K = 1$). Denote by $d = d_{ps}/d_{ss}$ the relative distance between the PU and the SU, where d_{ps} is the distance between the PU's receiver and the SU's transmitter, and d_{ss} is the distance between the SU's receiver and transmitter. Then, the relative path loss is given by d^α with $\alpha = 3$. Given the normalized noise, SNR is represented by P^{ave}.

SISO frequency-selective channels. As mentioned in Section V-A and shown in Appendix C, the robust power allocation problem (50) can be solved through the decomposition method. In Fig. 2, we show the iteration of the projection subgradient method for a randomly chosen channel realization, where the convergence to the optimal solution is observed. In Fig. 3, we depict the average rates and worst-case sum interference of the robust, non-robust, and perfect-CSI power allocation. By perfect CSI we mean that the transmitter knows the worst channel error for the non-robust strategy so that the interference constraint can be met exactly. The average is taken over $\boldsymbol{g}$ and $\boldsymbol{h}$ whose elements are randomly generated according to zero-mean, unit-variance, i.i.d. Gaussian distributions. One can see that the non-robust power allocation achieves the highest rate at the cost of violating the interference limit, which becomes severe at high SNR or when channel uncertainty increases. The perfect-CSI power allocation satisfies exactly the interference constraint with a lower rate, but perfect SU-to-PU CSI is seldom available in practice. The robust power allocation can provide an acceptable rate for the SU, and meanwhile never violates the interference constraint even with imperfect SU-to-PU CSI.

MIMO channels. In Fig 4, we show, respectively, the average rates and worst-case sum interference of different precoding techniques in a MIMO CR system. The average is taken over $\hat{\mathbf{G}}$ and $\mathbf{H}$, whose elements are randomly generated according to zero-mean, unit-variance, i.i.d. Gaussian distributions. From Fig. 4, one can also observe that, while the non-robust MIMO precoding may generate interference dramatically higher than the given threshold even for a small amount of uncertainty, the PU's communications can be efficaciously protected by the SU using the robust MIMO precoding that always complies with the interference restriction.

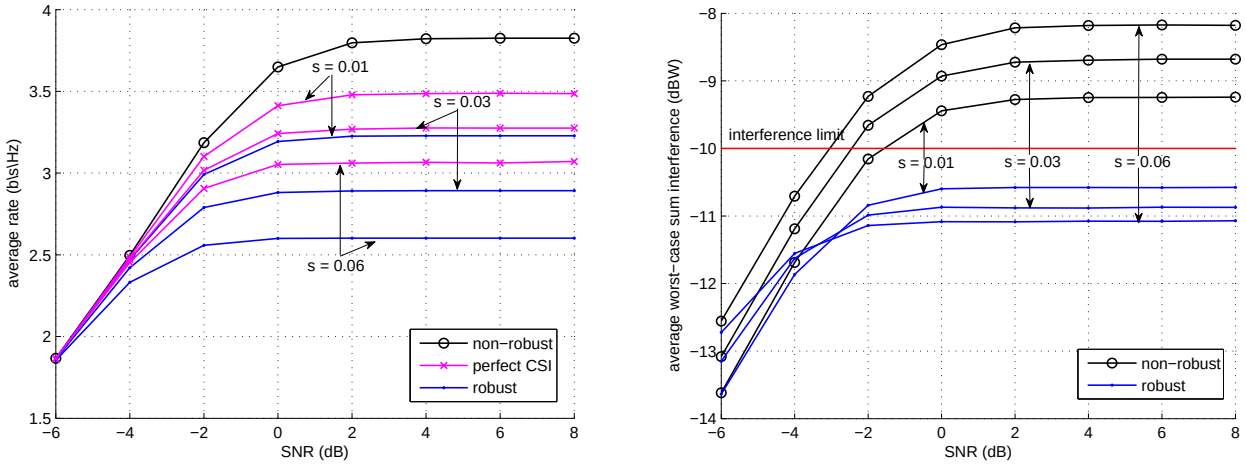
[3]Roughly speaking, a feasible strategy profile is a dominant strategy for a player if, *regardless of the strategies of the other players*, the strategy globally maximizes the payoff function of the player.

[4]Since the interference constraints are convex in errors, the maximum must be achieved on the boundary of the convex uncertainty region.



(a) (b)

Fig. 3. Comparison of the robust and non-robust power allocation for a single SU in SISO frequency-selective channels; $N = 16$, $I^{\text{ave}} = -10\text{dB}$, $d = 3$. (a) Average information rate. (b) Average worst-case sum interference.



(a) (b)

Fig. 4. Comparison of the robust and non-robust precoding for a single SU in MIMO channels; $M = N = 4$, $I^{\text{ave}} = -10\text{dB}$, $d = 3$. (a) Average information rate. (b) Average worst-case sum interference.

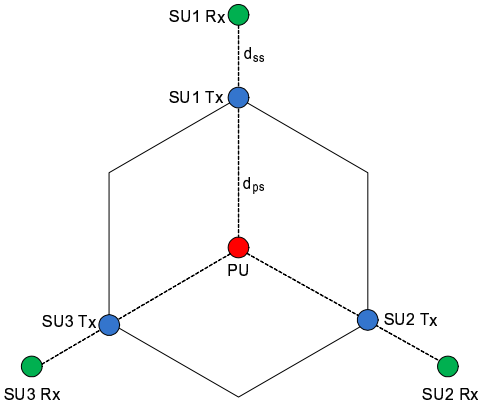


Fig. 5. Topology of a CR system with 3 SUs and 1 PU.

B. Multi-User Scenario

We consider a symmetric topology of a CR system as depicted in Fig 5, which includes one PU as the base station at the center of a hexagonal cell, and 3 SUs at the vertices of the hexagon with the same distance to their receivers. The relative path loss between the SUs and the PU is given by d^α, where $d = d_{ps}/d_{ss}$ is the relative distance and $\alpha = 3$. For simplicity, we impose to all SUs a common power budget ($P_q^{\text{ave}} = P^{\text{ave}}$, $\forall q$) and a common peak interference limit ($I_{qk}^{\text{peak}}, I_{qk}^{\text{peak}}(n) = I^{\text{peak}}$ and $I_{qk}^{\text{ave}}, I_{qk}^{\text{ave}}(n) = \infty$, $\forall n, q, k$). As a result, SNR of each SU is represented by P^{ave}.

SISO frequency-selective channels. To illustrate how the channel uncertainty affects the game, we show in Fig. 6 the probability of the condition $\rho(\mathbf{B}) < 1$ versus the relative distance d with randomly generated $\boldsymbol{h}_{qr}$ and $\hat{\boldsymbol{g}}_{qk}$, whose elements follow zero-mean, unit-variance, i.i.d. Gaussian distributions. Fig. 6 verifies our conclusion that the more uncertainty on SU-to-PU CSI (i.e., bigger s), the more likely the distributed algorithms converge to the unique NE of the game. Then, in Fig. 7, we demonstrate the iterative procedure of the sequential SISO RICA for $\hat{\mathcal{G}}_{\text{siso}}$ and $\mathcal{G}_{\text{siso}}$ over a randomly chosen channel profile (satisfying the condition $\rho(\mathbf{B}) < 1$) with the initial

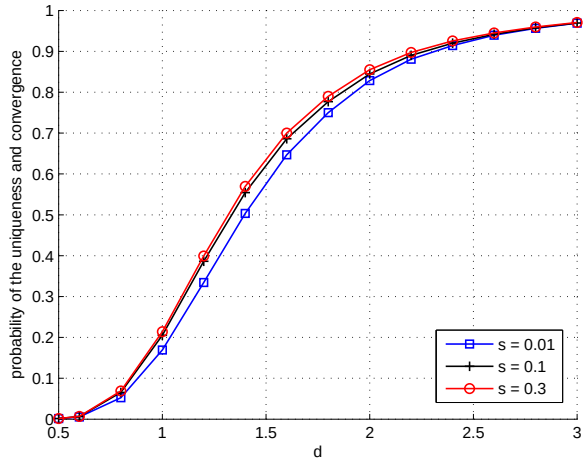


Fig. 6. Probability that the uniqueness and convergence conditions in Theorems 1 and 2 are satisfied; $Q = 3$, $K = 1$, $N = 8$, $I^{\text{ave}} = 0.1(-10\text{dB})$, $d = 2$, SNR $= 3$dB.

point from zero. The sequential updating is a special case of our proposed asynchoronous algorithm, where the SUs are scheduled in round bin. Upon the convergence of both the robust and non-robust designs, during the whole iteration, the robust strategy never breaks the interference limit, which is, however, violated by the non-robust strategy of each SU. Note that, in this case, the worst-case peak interference for a given strategy can be found in a simple and closed form as $\max_n \left(|\hat{g}_{qk}(n)| + \varepsilon_{qk}\right)^2 p_q(n)$ for each SU q.

MIMO channels. In Fig. 8, we demonstrate the iterative procedure of the simultaneous MIMO RICA, where the SUs simultaneously update their strategies at each iteration, for $\hat{\mathcal{G}}_{\text{mimo}}$ and $\mathcal{G}_{\text{mimo}}$ over a randomly chosen channel profile (satisfying the condition $\rho(\mathbf{D}_\alpha^{-1}\mathbf{B}_\beta) < 1$) with the initial point also from zero. Both the robust and non-robust designs converge very fast to the (unique) NEs of $\hat{\mathcal{G}}_{\text{mimo}}$ and $\mathcal{G}_{\text{mimo}}$, respectively. Similarly, Fig. 8 indicates that, at the cost of moderate loss of SUs' rates (compared to the non-robust strategy), the robust strategy can guarantee an interference level for any channel (error) within the uncertainty region, which is exactly our design goal in this paper.

VII. Conclusion

In this paper, we considered the robust design of a CR system, composed of multiple primary and secondary users, with imperfect SU-to-PU CSI. The resource allocation problem of the SU network has been formulated as a noncooperative game, in which each SU maximizes his information rate, subject to power constraints and robust interference constraints imposed by the PUs. We have analyzed the robust games using the advanced theory of VIs, and proposed totally asynchronous distributed algorithms along with their convergence properties. Furthermore, we proposed efficient methods to compute the best-response–the robust transmit strategy–of each SU.

Appendix A
Proof of Theorem 3

Before proving the theorem, we need some intermediate definitions and results on the VI theory, as given next.

Definition of the VI problem. Given a subset $\mathcal{K}$ of the Euclidean n-dimensional space $\mathbb{R}^n$ and a mapping $\mathbf{F} : \mathcal{K} \mapsto \mathbb{R}^n$, a VI problem, denoted by VI$(\mathcal{K}, \mathbf{F})$, is to find a vector $\mathbf{x}^\star \in \mathcal{K}$ such that

$$(\mathbf{x} - \mathbf{x}^\star)^T \mathbf{F}(\mathbf{x}^\star) \geq 0, \quad \forall \mathbf{x} \in \mathcal{K}. \tag{64}$$

Existence and uniqueness results. The theory and solution methods for various kinds of VIs are well developed. Here, we recall some of the basic conditions for the existence and uniqueness of the solution to a VI, as they are used in the paper. A classical existence result follows from [38, Corollary 2.2.5]: VI$(\mathcal{K}, \mathbf{F})$ is solvable if i) $\mathcal{K}$ is a nonempty, convex, and compact subset of a finite-dimensional Euclidean space; and ii) $\mathbf{F}$ is a continuous mapping. As far as the uniqueness of the solution, we refer to the following condition [38, Proposition 3.5.10]: Suppose that i) the set $\mathcal{K} \subseteq \mathbb{R}^n$ has a Cartesian structure, i.e., $\mathcal{K} = \prod_{q=1}^Q \mathcal{K}_q$ with $\mathcal{K}_q \subseteq \mathbb{R}^{n_q}$ and $n = \sum_{q=1}^Q n_q$, and each $\mathcal{K}_q$ is closed and convex; and ii) $\mathbf{F}(\mathbf{x}) \triangleq (\mathbf{F}_q(\mathbf{x}))_{q=1}^Q$ is a continuous uniformly-P function, i.e., there exists a constraint c_{up} such that for any $\mathbf{x} \triangleq (\mathbf{x}_q)_{q=1}^Q$ and $\mathbf{x}' \triangleq (\mathbf{x}'_q)_{q=1}^Q$ in $\mathcal{K}$

$$\max_q (\mathbf{x}_q - \mathbf{x}'_q)^T (\mathbf{F}_q(\mathbf{x}) - \mathbf{F}_q(\mathbf{x}')) \geq c_{\text{up}} \left\|\mathbf{x}_q - \mathbf{x}'_q\right\|^2. \tag{65}$$

Then, VI$(\mathcal{K}, \mathbf{F})$ has a unique solution.

Reformulation of the games $\hat{\breve{\mathcal{G}}}_{\text{mimo}}$ and $\breve{\mathcal{G}}_{\text{mimo}}$ as a VI problem. We focus only on the game $\breve{\mathcal{G}}_{\text{mimo}}$ without loss of generality. By convexity and the first-order (necessary and sufficient) optimality conditions [42] of each optimization problem (43), we infer that a strategy profile $\breve{\mathbf{Q}}^\star$ is a NE of $\breve{\mathcal{G}}_{\text{mimo}}$ if and only if

$$(\breve{\mathbf{Q}}_q - \breve{\mathbf{Q}}_q^\star)^T \bullet \nabla_{\breve{\mathbf{Q}}_q} \breve{R}_q(\breve{\mathbf{Q}}^\star) \leq 0, \ \forall \breve{\mathbf{Q}}_q \in \breve{\mathcal{Q}}_q, \ \forall q \tag{66}$$

where $\mathbf{A} \bullet \mathbf{B} \triangleq \text{Tr}(\mathbf{A}^T\mathbf{B}) = \text{vec}(\mathbf{A})^T \text{vec}(\mathbf{B})$. Define $\breve{\mathcal{Q}} \triangleq \breve{\mathcal{Q}}_1 \times \cdots \times \breve{\mathcal{Q}}_Q$ and $\mathbf{F}(\breve{\mathbf{Q}}) \triangleq (\mathbf{F}_q(\breve{\mathbf{Q}}))_{q=1}^Q$ where

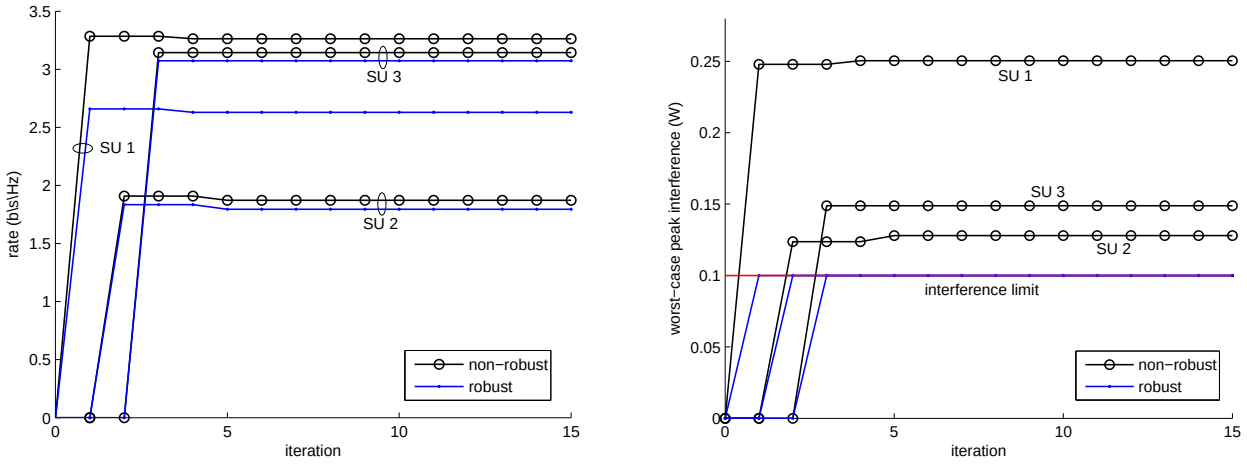
$$\mathbf{F}_q(\breve{\mathbf{Q}}) \triangleq -\nabla_{\breve{\mathbf{Q}}_q} \breve{R}_q(\breve{\mathbf{Q}}_q, \breve{\mathbf{Q}}_{-q}) = -\breve{\mathbf{H}}_{qq}^T \breve{\mathbf{Z}}_q^{-1}(\breve{\mathbf{Q}}) \breve{\mathbf{H}}_{qq} \tag{67}$$

with $\breve{\mathbf{Z}}_q(\breve{\mathbf{Q}}) \triangleq \breve{\mathbf{R}}_{n_q} + \sum_{r=1}^Q \breve{\mathbf{H}}_{rq} \breve{\mathbf{Q}}_r \breve{\mathbf{H}}_{rq}^T$. Then, $\breve{\mathbf{Q}}^\star$ is a NE of $\breve{\mathcal{G}}_{\text{mimo}}$ if and only if $\breve{\mathbf{Q}}^\star$ is a solution to VI$(\breve{\mathcal{Q}}, \mathbf{F})$. Therefore, we can resort to the VI theory to study $\breve{\mathcal{G}}_{\text{mimo}}$, or equivalently $\hat{\breve{\mathcal{G}}}_{\text{mimo}}$.

Proof of Theorem 3. The existence of a NE of $\breve{\mathcal{G}}_{\text{mimo}}$ (or $\hat{\breve{\mathcal{G}}}_{\text{mimo}}$) comes readily from the existence of a solution to VI$(\breve{\mathcal{Q}}, \mathbf{F})$, since the function $\mathbf{F}$ is continuous and the set $\breve{\mathcal{Q}}$ is compact and convex. Regarding the uniqueness, we show next that, under the conditions in Theorem 3 (b), $\mathbf{F}$ is a uniformly P-function on $\breve{\mathcal{Q}}$, implying the uniqueness of the solution of VI$(\breve{\mathcal{Q}}, \mathbf{F})$. To this end, we need the following definitions.

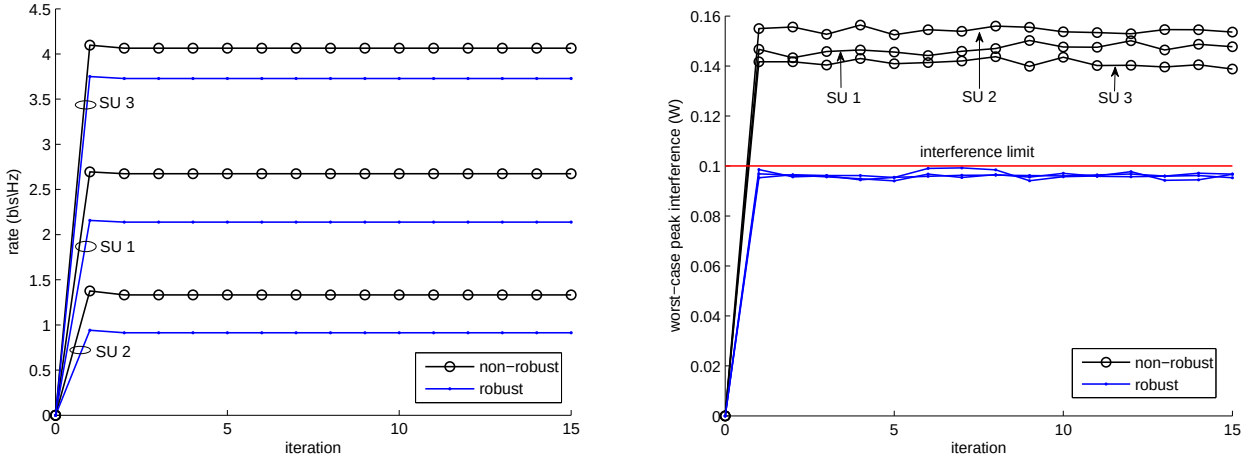
The Jacobian matrix of $\mathbf{F}_q$ evaluated at $\breve{\mathbf{Q}}$ is denoted by $\text{J}\mathbf{F}_q(\breve{\mathbf{Q}})$ and is given by [20]

$$\begin{aligned} \text{J}\mathbf{F}_q(\breve{\mathbf{Q}}) &\triangleq \frac{\partial \text{vec}\mathbf{F}_q(\breve{\mathbf{Q}})}{\partial \text{vec}(\breve{\mathbf{Q}})^T} = \left(\frac{\partial \text{vec}\mathbf{F}_q(\breve{\mathbf{Q}})}{\partial \text{vec}(\breve{\mathbf{Q}}_r)^T}\right)_{r=1}^Q \\ &= \left(\mathbf{\Gamma}_{rq}(\breve{\mathbf{Q}}) \otimes \mathbf{\Gamma}_{rq}(\breve{\mathbf{Q}})\right)_{r=1}^Q \\ &\triangleq \left(\text{J}_r\mathbf{F}_q(\breve{\mathbf{Q}})\right)_{r=1}^Q. \end{aligned} \tag{68}$$



(a) (b)

Fig. 7. Comparison of the robust game $\hat{\mathcal{G}}_{\text{siso}}$ and non-robust game $\mathcal{G}_{\text{siso}}$ using the sequential SISO RICA; $N = 8$, $I^{\text{ave}} = 0.1(-10\text{dB})$, $d = 2$, $\text{SNR} = 3\text{dB}$, $s = 0.04$. (a) Iteration of the SUs' rates. (b) Iteration of the SUs' worst-case peak interference.



(a) (b)

Fig. 8. Comparison of the robust game $\hat{\mathcal{G}}_{\text{mimo}}$ and non-robust game $\mathcal{G}_{\text{mimo}}$ using the simultaneous MIMO RICA; $M = N = 2$, $I^{\text{ave}} = 0.1(-10\text{dB})$, $d = 2.5$, $\text{SNR} = 3\text{dB}$, $s = 0.04$. (a) Iteration of the SUs' rates. (b) Iteration of the SUs' worst-case peak interference.

where $\mathbf{\Gamma}_{rq}(\breve{\mathbf{Q}}) \triangleq \breve{\mathbf{H}}_{qq}^T \breve{\mathbf{Z}}_q^{-1}(\breve{\mathbf{Q}}) \breve{\mathbf{H}}_{rq}$ and $\breve{\mathbf{Z}}_q(\breve{\mathbf{Q}}) \triangleq \breve{\mathbf{R}}_{n_q} + \sum_{r=1}^{Q} \breve{\mathbf{H}}_{rq} \breve{\mathbf{Q}}_r \breve{\mathbf{H}}_{rq}^T$. We also introduce

$$\alpha_q(\breve{\mathbf{Q}}) \triangleq \lambda_{\min}(\mathrm{J}_q \mathbf{F}_q(\breve{\mathbf{Q}})) = \lambda_{\min}^2(\mathbf{\Gamma}_{qq}(\breve{\mathbf{Q}})) \tag{69}$$

$$\beta_{rq}(\breve{\mathbf{Q}}) \triangleq \left\| \mathrm{J}_r \mathbf{F}_q(\breve{\mathbf{Q}}) \right\|_2 = \lambda_{\max}(\mathbf{\Gamma}_{rq}^T(\breve{\mathbf{Q}}) \mathbf{\Gamma}_{rq}(\breve{\mathbf{Q}})). \tag{70}$$

Using the property that a complex matrix $\mathbf{X}$ and its real counterpart $\breve{\mathbf{X}}$ have the same eigenvalues and the following inequalities: $\lambda_{\max}(\breve{\mathbf{Q}}_q) \leq \min\{P_q^{\text{ave}}, P_q^{\text{peak}}\}$, $\lambda_{\max}(\breve{\mathbf{Q}}_q) \leq \lambda_{\max}(\breve{\mathbf{T}}_{qk}) \min\{\mu_{qk}, \eta_{qk}\}$, $\lambda_{\max}(\breve{\mathbf{G}}_{qk} \breve{\mathbf{Q}}_q \breve{\mathbf{G}}_{qk}^T) \leq \min\{I_{qk}^{\text{ave}}, I_{qk}^{\text{peak}}\}$, $\mu_{qk} \varepsilon_{qk}^2 \leq I_{qk}^{\text{ave}}$, $\eta_{qk} \varepsilon_{qk}^2 \leq I_{qk}^{\text{peak}}$, $\forall k$, one can see that $\breve{\mathbf{Q}}_q \preceq \gamma_q \mathbf{I}$ for any $\breve{\mathbf{Q}}_q \in \breve{\mathcal{Q}}_q$, where γ_q is defined in (47). This implies

$$\alpha_q(\breve{\mathbf{Q}}) = \lambda_{\min}^2(\mathbf{\Gamma}_{qq}(\breve{\mathbf{Q}})) \geq \lambda_{\min}^2(\mathbf{\Gamma}_{qq}((\gamma_q \mathbf{I})_{q=1}^Q)) = \alpha_q^{\min} \tag{71}$$

where $\alpha_q^{\min}$ are defined in (45). We also need an upper bound of $\beta_{rq}(\breve{\mathbf{Q}})$:

$$\begin{aligned} \beta_{rq}(\breve{\mathbf{Q}}) &= \lambda_{\max}\left(\breve{\mathbf{H}}_{rq}^T \breve{\mathbf{Z}}_q^{-1}(\breve{\mathbf{Q}}) \breve{\mathbf{H}}_{qq} \breve{\mathbf{H}}_{qq}^T \breve{\mathbf{Z}}_q^{-1}(\breve{\mathbf{Q}}) \breve{\mathbf{H}}_{rq} \right) \\ &\leq \lambda_{\max}\left(\breve{\mathbf{H}}_{rq} \breve{\mathbf{H}}_{rq}^T \right) \lambda_{\max}\left(\breve{\mathbf{H}}_{qq}^T \breve{\mathbf{Z}}_q^{-2}(\breve{\mathbf{Q}}) \breve{\mathbf{H}}_{qq} \right) \\ &\leq \lambda_{\max}\left(\breve{\mathbf{H}}_{rq} \breve{\mathbf{H}}_{rq}^T \right) \lambda_{\max}\left(\breve{\mathbf{H}}_{qq}^T \breve{\mathbf{R}}_{n_q}^{-2} \breve{\mathbf{H}}_{qq} \right) \\ &= \beta_{rq}^{\max} \end{aligned} \tag{72}$$

where the first inequality follows from $\breve{\mathbf{R}}_{n_q} \preceq \mathbf{Z}_q(\breve{\mathbf{Q}})$, and $\beta_{rq}^{\max}$ is defined in (46).

Now we are ready to study the uniformly P-property of $\mathbf{F}$. Given two points $\breve{\mathbf{Q}}^{(1)} \neq \breve{\mathbf{Q}}^{(2)} \in \breve{\mathcal{Q}}$, we define for each q a real scalar function $\phi_q : [0,1] \ni v \mapsto \mathbb{R}$ as

$$\phi_q(v) \triangleq \left(\breve{\mathbf{Q}}_q^{(1)} - \breve{\mathbf{Q}}_q^{(2)} \right) \bullet \mathbf{F}_q \left(v \breve{\mathbf{Q}}^{(1)} + (1-v) \breve{\mathbf{Q}}^{(2)} \right) \tag{73}$$

Note that $\phi_q(v)$ is continuously differentiable on $(0,1)$. We

can then apply the mean-value theorem and write

$$\begin{aligned}&\left(\breve{\mathbf{Q}}_q^{(1)}-\breve{\mathbf{Q}}_q^{(2)}\right)\bullet\left(\mathbf{F}_q(\breve{\mathbf{Q}}^{(1)})-\mathbf{F}_q(\breve{\mathbf{Q}}^{(2)})\right)\\ =\;&\phi_q(1)-\phi_q(0)=\phi_q'(\bar{v})\end{aligned}\tag{74}$$

for some $\bar{v}\in(0,1)$. Let $\mathbf{X}_{\bar{v}}\triangleq\bar{v}\breve{\mathbf{Q}}^{(1)}+(1-\bar{v})\breve{\mathbf{Q}}^{(2)}$. Then, $\phi_q'(\bar{v})$ is given by

$$\begin{aligned}&\phi_q'(\bar{v})\\ =\;&\mathrm{vec}\left(\breve{\mathbf{Q}}_q^{(1)}-\breve{\mathbf{Q}}_q^{(2)}\right)^T\mathrm{J}\mathbf{F}_q\left(\mathbf{X}_{\bar{v}}\right)\mathrm{vec}\left(\breve{\mathbf{Q}}^{(1)}-\breve{\mathbf{Q}}^{(2)}\right)\\ =\;&\mathrm{vec}\left(\breve{\mathbf{Q}}_q^{(1)}-\breve{\mathbf{Q}}_q^{(2)}\right)^T\sum_{r=1}^{Q}\mathrm{J}_r\mathbf{F}_q\left(\mathbf{X}_{\bar{v}}\right)\cdot\\ &\mathrm{vec}\left(\breve{\mathbf{Q}}_r^{(1)}-\breve{\mathbf{Q}}_r^{(2)}\right)\end{aligned}\tag{75}$$

Introducing $\mathbf{d}_q\triangleq\mathrm{vec}\left(\breve{\mathbf{Q}}_q^{(1)}-\breve{\mathbf{Q}}_q^{(2)}\right)$, it follows that

$$\begin{aligned}\phi_q'(\bar{v})&=\mathbf{d}_q^T\sum_{r=1}^{Q}\mathrm{J}_r\mathbf{F}_q\left(\mathbf{X}_{\bar{v}}\right)\mathbf{d}_r\\ &\geq\mathbf{d}_q^T\mathrm{J}_q\mathbf{F}_q\left(\mathbf{X}_{\bar{v}}\right)\mathbf{d}_q-\sum_{r\neq q}\left|\mathbf{d}_q^T\mathrm{J}_r\mathbf{F}_q\left(\mathbf{X}_{\bar{v}}\right)\mathbf{d}_r\right|&(76)\\ &\geq\left\|\mathbf{d}_q\right\|^2\lambda_{\min}(\mathrm{J}_q\mathbf{F}_q\left(\mathbf{X}_{\bar{v}}\right))\\ &\quad-\left\|\mathbf{d}_q\right\|\sum_{r\neq q}\left\|\mathrm{J}_r\mathbf{F}_q\left(\mathbf{X}_{\bar{v}}\right)\right\|_2\left\|\mathbf{d}_r\right\|&(77)\\ &=\left\|\mathbf{d}_q\right\|^2\alpha_q(\mathbf{X}_{\bar{v}})-\left\|\mathbf{d}_q\right\|\sum_{r\neq q}\beta_{rq}(\mathbf{X}_{\bar{v}})\left\|\mathbf{d}_r\right\|&(78)\\ &\geq\left\|\mathbf{d}_q\right\|^2\alpha_q^{\min}-\left\|\mathbf{d}_q\right\|\sum_{r\neq q}\beta_{rq}^{\max}\left\|\mathbf{d}_r\right\|&(79)\end{aligned}$$

where (77) follows from the Cauchy-Schwarz inequality and $\|\mathbf{Ab}\|\leq\|\mathbf{A}\|_2\|\mathbf{b}\|$; (79) follows from the fact that $\mathbf{X}_{\bar{v}}$, as a convex combination of $\breve{\mathbf{Q}}_q^{(1)}$ and $\breve{\mathbf{Q}}_q^{(2)}$, belongs to $\breve{\mathcal{Q}}_q$.

Introducing $\mathbf{s}\triangleq(s_q)_{q=1}^Q$ with $s_q\triangleq\|\mathbf{d}_q\|$, and $\boldsymbol{\Upsilon}\triangleq\mathbf{D}_\alpha-\mathbf{B}_\beta$ where $\mathbf{D}_\alpha$ and $\mathbf{B}_\beta$ are defined in (44). from (74), (79), we have

$$\left(\breve{\mathbf{Q}}_q^{(1)}-\breve{\mathbf{Q}}_q^{(2)}\right)\bullet\left(\mathbf{F}_q(\breve{\mathbf{Q}}^{(1)})-\mathbf{F}_q(\breve{\mathbf{Q}}^{(2)})\right)\geq s_q\left(\boldsymbol{\Upsilon}\mathbf{s}\right)_q.\tag{80}$$

If $\boldsymbol{\Upsilon}$ is a P-matrix[5], then there exists a constant $c_{\mathrm{up}}(\boldsymbol{\Upsilon})$ such that [24]

$$\max_{q\in\Omega}x_q\left(\boldsymbol{\Upsilon}\mathbf{x}\right)_q\geq c_{\mathrm{up}}(\boldsymbol{\Upsilon})\left\|\mathbf{x}\right\|^2\tag{81}$$

holds for an arbitrary vector $\mathbf{x}\in\mathbb{R}^Q$. Therefore, under the P-property of $\boldsymbol{\Upsilon}$, we have

$$\begin{aligned}&\max_{q\in\Omega}\left(\breve{\mathbf{Q}}_q^{(1)}-\breve{\mathbf{Q}}_q^{(2)}\right)\bullet\left(\mathbf{F}_q(\breve{\mathbf{Q}}^{(1)})-\mathbf{F}_q(\breve{\mathbf{Q}}^{(2)})\right)\\ \geq\;&\max_{q\in\Omega}s_q\left(\boldsymbol{\Upsilon}\mathbf{s}\right)_q\geq c_{\mathrm{up}}(\boldsymbol{\Upsilon})\left\|\breve{\mathbf{Q}}_q^{(1)}-\breve{\mathbf{Q}}_q^{(2)}\right\|_F^2\end{aligned}\tag{82}$$

for $\forall\breve{\mathbf{Q}}^{(1)},\breve{\mathbf{Q}}^{(2)}\in\breve{\mathcal{Q}}$, implying the uniformly P-property of $\mathbf{F}$. Note that a necessary condition for $\boldsymbol{\Upsilon}$ to be a P-matrix is that $\alpha_q^{\min}>0$, $\forall q$, which is equivalent to the requirement that $\mathbf{H}_{q,q}$ is full column-rank for all q. Invoking [46, Lemma 5.13.14] and following the same steps as in [24, Appendix B], it is not difficult to show the equivalence: $\boldsymbol{\Upsilon}$ is a P-matrix $\Leftrightarrow\rho(\mathbf{D}_\alpha^{-1}\mathbf{B}_\beta)<1$, which completes the proof.

[5]A matrix $\mathbf{M}\in\mathbb{R}^{n\times n}$ is called called P-matrix if its every principal minor is positive.

Appendix B
Proof of Theorem 4

To prove the convergence of the asynchronous RICA, we can follow the similar steps as in the proof of [20, Theorem 12], according to which it is enough to prove the convergence of the simultaneous RICA. In the simultaneous RICA, all SUs update their strategies at each iteration (thus $\tau_r^q(m)=m$), and at the m^{th} iteration SU q chooses his best response based on the strategies of the others $\mathbf{Q}_{-q}^{(m)}\triangleq\left(\mathbf{Q}_1^{(m)},\ldots,\mathbf{Q}_{q-1}^{(m)},\mathbf{Q}_{q+1}^{(m)},\ldots,\mathbf{Q}_Q^{(m)}\right)$.

Given the equivalence of $\breve{\mathcal{G}}_{\mathrm{mimo}}$ and $\hat{\mathcal{G}}_{\mathrm{mimo}}$, we focus on $\breve{\mathcal{G}}_{\mathrm{mimo}}$. For each SU q, $\breve{\mathbf{Q}}_q^{(m+1)}$ is given by the solution to (43) with $\breve{\mathbf{Q}}_{-q}=\breve{\mathbf{Q}}_{-q}^{(m)}$, and satisfies the first-order optimality condition

$$\left(\breve{\mathbf{Q}}_q-\breve{\mathbf{Q}}_q^{(m+1)}\right)\bullet\mathbf{F}_q\left(\breve{\mathbf{Q}}_q^{(m+1)},\breve{\mathbf{Q}}_q^{(m)}\right)\geq0,\;\forall\breve{\mathbf{Q}}_q\in\breve{\mathcal{Q}}_q.\tag{83}$$

Adding (66) evaluated at $\breve{\mathbf{Q}}_q^{(m+1)}$ and (83) evaluated at the NE $\breve{\mathbf{Q}}_q^\star$, we have

$$\begin{aligned}&\left(\breve{\mathbf{Q}}_q^{(m+1)}-\breve{\mathbf{Q}}_q^\star\right)\bullet\left(\mathbf{F}_q(\breve{\mathbf{Q}}_q^\star,\breve{\mathbf{Q}}_{-q}^\star)-\mathbf{F}_q(\breve{\mathbf{Q}}_q^{(m+1)},\breve{\mathbf{Q}}_q^{(m)})\right)\\ &\geq0,\;\forall q.\end{aligned}\tag{84}$$

Now we can apply the mean-value theorem to (84). Introducing $\mathbf{Y}_{\bar{v}}\triangleq v(\breve{\mathbf{Q}}_q^\star,\breve{\mathbf{Q}}_{-q}^\star)+(1-v)(\breve{\mathbf{Q}}_q^{(m+1)},\breve{\mathbf{Q}}_q^{(m)})$ and following the steps (73)-(75), we have for some $\bar{v}\in(0,1)$

$$\begin{aligned}0\;\leq\;&\mathrm{vec}\left(\breve{\mathbf{Q}}_q^{(m+1)}-\breve{\mathbf{Q}}_q^\star\right)^T\mathrm{J}\mathbf{F}_q\left(\mathbf{Y}_{\bar{v}}\right)\cdot\\ &\mathrm{vec}\left((\breve{\mathbf{Q}}_q^\star,\breve{\mathbf{Q}}_{-q}^\star)-(\breve{\mathbf{Q}}_q^{(m+1)},\breve{\mathbf{Q}}_q^{(m)})\right)\\ =\;&-(\mathbf{d}_q^{(m+1)})^T\mathrm{J}_q\mathbf{F}_q(\mathbf{Y}_{\bar{v}})\mathbf{d}_q^{(m+1)}\\ &-(\mathbf{d}_q^{(m+1)})^T\sum_{r\neq q}\mathrm{J}_r\mathbf{F}_q(\mathbf{Y}_{\bar{v}})\mathbf{d}_r^{(m)}\end{aligned}\tag{85}$$

where $\mathbf{d}_q^{(m+1)}\triangleq\mathrm{vec}\left(\breve{\mathbf{Q}}_q^{(m+1)}-\breve{\mathbf{Q}}_q^\star\right)$ and $\mathbf{d}_q^{(m)}\triangleq\mathrm{vec}\left(\breve{\mathbf{Q}}_q^{(m)}-\breve{\mathbf{Q}}_q^\star\right)$. Recalling the definitions of $\alpha_q(\mathbf{Y}_{\bar{v}})$ and $\beta_{rq}(\mathbf{Y}_{\bar{v}})$ in (69) and (70), respectively, and following the similar steps (76)-(79), we obtain from (85)

$$\begin{aligned}&\left\|\mathbf{d}_q^{(m+1)}\right\|\alpha_q^{\min}\\ \leq\;&\left\|\mathbf{d}_q^{(m+1)}\right\|\alpha_q(\mathbf{Y}_{\bar{v}})\leq\sum_{r\neq q}\beta_{rq}(\mathbf{Y}_{\bar{v}})\left\|\mathbf{d}_r^{(m)}\right\|\\ \leq\;&\sum_{r\neq q}\beta_{rq}^{\max}\left\|\mathbf{d}_r^{(m)}\right\|.\end{aligned}\tag{86}$$

Assuming that $\mathbf{H}_{q,q}$ is full column-rank, $\forall q$, then $\alpha_q^{\min}>0$, $\forall q$, then (86) implies

$$\left\|\mathbf{d}_q^{(m+1)}\right\|\leq\frac{1}{\alpha_q^{\min}}\sum_{r\neq q}\beta_{rq}^{\max}\left\|\mathbf{d}_r^{(m)}\right\|,\;\forall q.\tag{87}$$

Introducing $\mathbf{s}^{(m+1)}\triangleq(s_q^{(m+1)})_{q=1}^Q$ with $s_q^{(m+1)}\triangleq\left\|\mathbf{d}_q^{(m+1)}\right\|$, and recalling the definitions of $\mathbf{D}_\alpha$ and $\mathbf{B}_\beta$ in (44), (87) can

be rewritten as

$$\mathbf{s}^{(m+1)} \leq \mathbf{D}_{\alpha}^{-1}\mathbf{B}_{\beta}\mathbf{s}^{(m)}. \tag{88}$$

It follows that the error sequence $\{\mathbf{s}^{(m)}\}_{m=1}^{\infty}$, generated by the simultaneous RICA, converges to zero if

$$\rho\left(\mathbf{D}_{\alpha}^{-1}\mathbf{B}_{\beta}\right) < 1 \tag{89}$$

which thus guarantees the global convergence of Algorithm 2.

APPENDIX C
DECOMPOSITION METHOD FOR SINGLE-USER OPTIMIZATION IN $\hat{\mathcal{G}}_{\text{siso}}$

For simplicity, we suppress the subscript q for all notations, since the optimization problem (50) is structurally the same for all SUs. Define $a(n) \triangleq |h_{qq}(n)|^2/(\sigma_q^2(n) + \sum_{r\neq q}|h_{rq}(n)|^2 p_r(n))$, $b_0(n) \triangleq 1/P_q^{\text{ave}}$, $b_k(n) \triangleq |\hat{g}_{qk}(n)|^2/I_{qk}^{\text{ave}}$, $c_0(n) \triangleq p_q^{\text{peak}}(n)$, $c_k(n) \triangleq I_{qk}^{\text{peak}}(n)/\left(|\hat{g}_{qk}(n)| + \varepsilon_{qk}/\sqrt{w_{qk}(n)}\right)^2$, $c^{\min}(n) \triangleq \min_{k=0,\ldots,K} c_k(n)$, $\theta_k \triangleq \varepsilon_{qk}^2/I_{qk}^{\text{ave}}$, and

$$\varphi_{nk}(p(n),\mu_k) \triangleq \frac{\mu_k w_k(n) b_k(n) p(n)}{\mu_k w_k(n) - p(n)} \quad \forall n,k. \tag{90}$$

Consequently, the convex problem (50) can be expressed as

$$\begin{array}{ll} \underset{\mathbf{p},\boldsymbol{\mu}}{\text{maximize}} & \sum_{n=1}^{N}\log(1+a(n)p(n)) \\ \text{subject to} & 0 \leq p(n) \leq c^{\min}(n),\ \forall n \\ & p(n) \leq \mu_k w_k(n),\ \forall n,k \\ & \sum_{n=1}^{N} b_0(n)p(n) \leq 1 \\ & \sum_{n=1}^{N}\varphi_{nk}(p(n),\mu_k) + \mu_k\theta_k \leq 1,\ \forall k. \end{array} \tag{91}$$

This problem has a decomposable structure in terms of the variables $\{p(n)\}_{n=1}^{N}$, so one can exploit multilevel decomposition methods [8] to further simply it. In the following, we elaborate how to use the so-called dual-primal decomposition method to solve (91). Note that other decomposition methods, e.g., the primal-dual decomposition, are also applicable.

The third-level subproblem. To avoid meaningless results in the subproblems, we first introduce the following redundant constraints into (91):

$$\varphi_{nk}(p(n),\mu_k) + \mu_k\theta_k \leq 1, \quad \forall n,k \tag{92}$$

which do not change the optimal solution to (91). Then, for each n, at the lowest (third) level is the subproblem

$$\begin{array}{ll} \underset{p(n)}{\text{maximize}} & f_n(p(n),\boldsymbol{\mu},\boldsymbol{\lambda}) \\ \text{subject to} & 0 \leq p(n) \leq c^{\min}(n) \\ & p(n) \leq \mu_k w_k(n),\ \forall k \\ & \varphi_{nk}(p(n),\mu_k) + \mu_k\theta_k \leq 1,\ \forall k \end{array} \tag{93}$$

where

$$\begin{aligned} f_n(p(n),\boldsymbol{\mu},\boldsymbol{\lambda}) \triangleq\ & \log(1+a(n)p(n)) - \lambda_0 b_0(n)p(n) \\ & - \sum_{k=1}^{K}\lambda_k\varphi_{nk}(p(n),\mu_k) \end{aligned} \tag{94}$$

with given $\boldsymbol{\mu} \triangleq (\mu_k)_{k=1}^{K}$ and $\boldsymbol{\lambda} \triangleq (\lambda_k)_{k=0}^{K}$. The redundant constraints in (92) prevent the subproblem (93) to reach the meaningless point $p(n) = \mu_k w_k(n)$ (see the proof of Proposition 3). Hence, the feasible set of (93) is an interval $[0, p^{\max}(n)]$, where $p^{\max}(n) \triangleq \min\{c^{\min}(n), d^{\min}(n)\}$ and

$$d^{\min}(n) \triangleq \min_{k} \frac{\mu_k w_k(n)(1-\mu_k\theta_k)}{\mu_k w_k(n) b_k(n) + 1 - \mu_k\theta_k}. \tag{95}$$

Let $h_n(p(n)) \triangleq \partial f_n(p(n),\boldsymbol{\mu},\boldsymbol{\lambda})/\partial p(n)$. Then, we have

$$h_n(p(n)) = \frac{a(n)}{1+a(n)p(n)} - \lambda_0 b_0(n) - \sum_{k=1}^{K}\frac{\lambda_k\mu_k^2 w_k^2(n) b_k(n)}{(\mu_k w_k(n) - p(n))^2} \tag{96}$$

which is monotonically decreasing on $[0, p^{\max}(n)]$. Therefore, the solution to (93) is just the root of $h_n(p(n))$ on $[0, p^{\max}(n)]$, which can be efficiently found via the bisection method.

The second-level subproblem. Denote the solution to (93) by $p^{\star}(n)$, and the optimum value by $f_n^{\star}(\boldsymbol{\mu},\boldsymbol{\lambda}) \triangleq f_n(p^{\star}(n),\boldsymbol{\mu},\boldsymbol{\lambda})$ for given $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$. Then, at the middle (second) level is a maximization problem with the variable $\boldsymbol{\mu}$ for given $\boldsymbol{\lambda}$ as

$$\underset{\mathbf{0}\leq\boldsymbol{\mu}\leq\boldsymbol{\mu}^{\max}}{\text{maximize}} \ \sum_{n=1}^{N} f_n^{\star}(\boldsymbol{\mu},\boldsymbol{\lambda}) - \sum_{k=1}^{K}\lambda_k\mu_k\theta_k \tag{97}$$

where $\boldsymbol{\mu}^{\max} \triangleq (\mu_k^{\max})_{k=1}^{K}$ with $\mu_k^{\max} \triangleq 1/\theta_k$.

The first-level master problem. Denote the optimum value of (97) by $g(\boldsymbol{\lambda})$. Finally, at the top (first) level is a minimization problem with the variable $\boldsymbol{\lambda}$ as

$$\underset{\boldsymbol{\lambda}\geq\mathbf{0}}{\text{minimize}} \ g(\boldsymbol{\lambda}) + \sum_{k=0}^{K}\lambda_k. \tag{98}$$

To solve (97) and (98), we rely on the subgradient-based algorithms. The following proposition provides the subgradients[6] of the objectives in (97) and (98).

Proposition 5: (a) Given $\boldsymbol{\lambda}$ and $\mathbf{p}^{\star} \triangleq (p^{\star}(n))_{n=1}^{N}$, a subgradient of $f_n^{\star}(\boldsymbol{\mu},\boldsymbol{\lambda})$ with respect to $\boldsymbol{\mu}$ is

$$\mathbf{s}_n(\boldsymbol{\mu}) = \left(\frac{\lambda_k w_k(n) b_k(n)(p^{\star}(n))^2}{(\mu_k w_k(n) - p^{\star}(n))^2} - \alpha_k(n)\right)_{k=1}^{K}. \tag{99}$$

If $\varphi_{nk}(p^{\star}(n),\mu_k) + \mu_k\theta_k < 1$, $\forall k$, then $\alpha_k(n) = 0$, $\forall k$; if $\varphi_{nl}(p^{\star}(n),\mu_l) + \mu_l\theta_l = 1$ for some l, then

$$\alpha_k(n) = \begin{cases} \left(\theta_l - \frac{w_l(n)b_l(n)(p^{\star}(n))^2}{(\mu_l w_l(n) - p^{\star}(n))^2}\right)\beta_l(n), & \text{if } k = l \\ 0, & \text{if } k \neq l \end{cases} \tag{100}$$

where

$$\beta_l(n) = h_n(p^{\star}(n),\boldsymbol{\mu},\boldsymbol{\lambda})\frac{(\mu_l w_l(n) - p^{\star}(n))^2}{\mu_l^2 w_l^2(n) b_l(n)}. \tag{101}$$

(b) Given $\boldsymbol{\mu}^{\star}$ as the optimal solution to (97) and $\mathbf{p}^{\star}$, a subgradient of $g(\boldsymbol{\lambda})$ is

$$\begin{aligned} \mathbf{s}_g(\boldsymbol{\lambda}) \ =\ & \left(-\sum_{n=1}^{N} b_0(n)p^{\star}(n)\ , \right. \\ & \left. \left(-\sum_{n=1}^{N}\varphi_{nk}(p^{\star}(n),\mu_k^{\star}) - \mu_k^{\star}\theta_k\right)_{k=1}^{K}\right) \end{aligned} \tag{102}$$

[6]Given a convex function $f : \mathbb{R}^n \longmapsto \mathbb{R}$, a vector $\mathbf{s}$ is a subgradient of f at a point $\mathbf{x} \in \mathbb{R}^n$ if $f(\mathbf{z}) \geq f(\mathbf{x}) + \mathbf{s}^T(\mathbf{z}-\mathbf{x})$, $\forall \mathbf{z} \in \mathbb{R}^n$. If instead f is a concave function, then $\mathbf{s}$ is a subgradient of f at $\mathbf{x}$ if $-\mathbf{s}$ is a subgradient of the convex function $-f$ at $\mathbf{x}$.

Proof: The calculation of the above subgradients is based on [8, Lemma 1]. Due to the space limitation, we refer the interested readers to [8] for more details. ■

Given the subgradients in Proposition 5, one can choose different subgradient methods, with different tradeoffs between the complexity and the convergence rate, to find the optimal solutions to (97) and (98). For example, the cutting-plane method may converge faster but has relatively high complexity. Here, we briefly introduce the subgradient projection method that may converge slowly but is very simple. Specifically, consider a convex problem

$$\underset{\mathbf{x}\geq\mathbf{0}}{\text{minimize}} \ f(\mathbf{x}). \tag{103}$$

Given a subgradient $\mathbf{s}(\mathbf{x})$ of $f(\mathbf{x})$, $\mathbf{x}$ is iteratively updated according to

$$\mathbf{x}^{(t+1)} = \left[\mathbf{x}^{(t)} - \kappa^{(t)}\mathbf{s}(\mathbf{x}^{(t)})\right]_+ \tag{104}$$

where $\kappa^{(t)}$ is the step size. With the properly chosen step size, for example a diminishing step size $\kappa^{(t)} = \kappa^{(0)}(1+z)/(t+z)$, where $\kappa^{(0)} \in (0,1]$ is the initial step size and z is a fixed nonnegative number, the sequence $\{\mathbf{x}^{(t)}\}_{t=1}^{\infty}$ is guaranteed to converge to the optimal solution to (103). Note that, when $K = 1$, i.e., only one PU, the complicated cutting-plane method reduces to the simple bisection method.

REFERENCES

[1] FCC Spectrum Policy Task Force, "Report of the spectrum efficiency working group," Tech. Rep. 02-135, Nov. 2002.

[2] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Commun.*, vol. 6, no. 6, pp. 13–18, Aug. 1999.

[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[4] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[5] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inform. Theory*, vol. 53, no. 2, pp. 471–487, Feb. 2007.

[6] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 88–102, Feb. 2008.

[7] M. Chiang, C. W. Tan, D. P. Palomar, D. O. Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[8] D. P. Palomar and M. Chiang, "Alternative distributed algorithms for network utility maximization: Framework and applications," *IEEE Trans. Autom. Control*, vol. 52, no. 12, pp. 2254–2269, Dec. 2007.

[9] E. G. Larsson, E. A. Jorswieck, J. Lindblom, and R. Mochaourab, "Game theory and the flat fading Gaussian interference channel," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 18–27, Sep. 2009.

[10] A. Leshem and E. Zehavi, "Game theory and the frequency selective interference channel - a tutorial," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 28–40, Sep. 2009.

[11] Z. Luo and S. Zhang, "Spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–72, Feb. 2008.

[12] M. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA: MIT Press, 1994.

[13] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[14] Z. Luo and J.-S. Pang, "Analysis of iterative waterfilling algorithm for multiuser power control in digital subscriber lines," *EURASIP Jour. on Applied Signal Process.*, vol. 2006, pp. 1–10, Feb. 2006.

[15] G. Scutari, D. P. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory—part i: Nash equilibria," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1230–1249, Mar. 2008.

[16] ——, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory—part ii: Algorithms," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1250–1267, Mar. 2008.

[17] ——, "Asynchronous iterative water-filling for Gaussian frequency-selective interference channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 7, pp. 2868–2878, Jul. 2008.

[18] J.-S. Pang, G. Scutari, F. Facchinei, and C. Wang, "Distributed power allocation with rate constraints in Gaussian parallel interference channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3471–3489, Aug. 2008.

[19] G. Scutari, D. P. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: A unified view," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1089–1103, Sep. 2008.

[20] ——, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, Mar. 2009.

[21] ——, "Cognitive MIMO radio: Competitive optimality design based on subspace projections," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 46–59, Nov. 2008.

[22] G. Scutari, D. P. Palomar, J.-S. Pang, and F. Facchinei, "Flexible design of cognitive radio wireless systems: From game theory to variational inequality theory," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 107–123, Sep. 2009.

[23] G. Scutari and D. P. Palomar, "MIMO cognitive radio: A game-theoretical approach," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 761–780, Feb. 2010.

[24] J.-S. Pang, G. Scutari, D. P. Palomar, and F. Facchinei, "Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3251–3271, Jun. 2010.

[25] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 2, pp. 662–671, Mar. 2006.

[26] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.

[27] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[28] Y. C. Eldar and N. Merhav, "A competitive minimax approach to robust estimation of random parameters," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1931–1946, Jul. 2004.

[29] R. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.

[30] A. Pascual-Iserte, D. P. Palomar, A. I. Pérez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with partial channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 346–360, Jan. 2006.

[31] J. Wang and D. P. Palomar, "Worst-case robust MIMO transmission with imperfect channel knowledge," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3086–3100, Aug. 2009.

[32] ——, "Robust MSE precoding in MIMO channels with pre-fixed receivers," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5802–5818, Nov. 2010.

[33] J. Wang and M. Payaró, "On the robustness of transmit beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5933–5938, Nov. 2010.

[34] L. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor, "Robust cognitive beamforming with partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4143–4153, Aug. 2009.

[35] G. Zheng, K.-K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4871–4881, Dec. 2009.

[36] A. J. Anandkumar, A. Anandkumar, S. Lambotharan, and J. Chambers, "Robust rate maximization game under bounded channel uncertainty," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, USA, Mar. 2010.

[37] G. Scutari, D. P. Palomar, J.-S. Pang, and F. Facchinei, "Convex optimization, game theory, and variational inequality theory," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 35–49, May 2010.

[38] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. New York: Springer, 2003.

[39] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.

[40] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1991.
[41] R. Zhang, "On peak versus average interference power constraints for protecting primary users in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2112–2120, Apr. 2009.
[42] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.
[43] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, Studies in Applied Mathematics 13, 1994.
[44] Z. Luo, J. F. Sturm, and S. Zhang, "Multivariate nonnegative quadratic mappings," *SIAM J. Optim.*, vol. 14, no. 4, pp. 1140–1162, 2004.
[45] G. A. F. Seber, *A Matrix Handbook for Statisticians*. New Jersey, US: John Wiley & Sons, 2008.
[46] R. W. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*. Cambridge Academic Press, 1992.

Jiaheng Wang (S'08-M'10) received the Ph.D. degree in electrical engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2010. He received the B.E. and M.S. degrees from the Southeast University, Nanjing, China, in 2001 and 2006, respectively. He is currently a postdoctor with the Signal Processing Laboratory, ACCESS Linnaeus Center, KTH Royal Institute of Technology, Stockholm, Sweden. His research interests mainly include optimization in signal processing, wireless communications and networks.

Gesualdo Scutari (S'05-M'06) received the Electrical Engineering and Ph.D. degrees (both with honors) from the University of Rome "La Sapienza," Rome, Italy, in 2001 and 2004, respectively.

He is an Assistant Professor in the Department of Electrical Engineering at State University of New York (SUNY) at Buffalo, Buffalo, NY. He had previously held several research appointments, namely, at University of California at Berkeley, Berkeley, CA; Hong Kong University of Science and Technology, Hong Kong; University of Rome, "La Sapienza", Rome, Italy; University of Illinois at Urbana-Champaign, Urbana, IL.

He has participated in many European projects, namely: IST SATURN (on multiantenna systems), IST-ROMANTIK (on multihop systems), WINSOC (on wireless sensor networks), and SURFACE (on reconfigurable air interfaces for wideband multiantenna communication systems). His primary research interests include applications of convex optimization theory, game theory, and variational inequality theory to signal processing and communications; sensor networks; and distributed decisions.

Dr. Scutari received the 2006 Best Student Paper Award at the International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2006.

Daniel P. Palomar (S'99-M'03-SM'08) received the Electrical Engineering and Ph.D. degrees (both with honors) from the Technical University of Catalonia (UPC), Barcelona, Spain, in 1998 and 2003, respectively.

He is an Associate Professor in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology (HKUST), Hong Kong, which he joined in 2006. He had previously held several research appointments, namely, at King's College London (KCL), London, UK; Technical University of Catalonia (UPC), Barcelona; Stanford University, Stanford, CA; Telecommunications Technological Center of Catalonia (CTTC), Barcelona; Royal Institute of Technology (KTH), Stockholm, Sweden; University of Rome "La Sapienza", Rome, Italy; and Princeton University, Princeton, NJ. His current research interests include applications of convex optimization theory, game theory, and variational inequality theory to signal processing and communications.

Dr. Palomar is an Associate Editor of IEEE Transactions on Information Theory and has been an Associate Editor of IEEE Transactions on Signal Processing, a Guest Editor of the IEEE Signal Processing Magazine 2010 Special Issue on "Convex Optimization for Signal Processing," a Guest Editor of the IEEE Journal on Selected Areas in Communications 2008 Special Issue on "Game Theory in Communication Systems," the Lead Guest Editor of the IEEE Journal on Selected Areas in Communications 2007 Special Issue on "Optimization of MIMO Transceivers for Realistic Communication Networks." He serves on the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications (SPCOM). He was the General Co-Chair of the 2009 IEEE Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP).

He is a recipient of a 2004/06 Fulbright Research Fellowship; the 2004 Young Author Best Paper Award by the IEEE Signal Processing Society; the 2002/03 best Ph.D. prize in Information Technologies and Communications by the Technical University of Catalonia (UPC); the 2002/03 Rosina Ribalta first prize for the Best Doctoral Thesis in Information Technologies and Communications by the Epson Foundation; and the 2004 prize for the best Doctoral Thesis in Advanced Mobile Communications by the Vodafone Foundation and COIT.